
EX LIBRIS
AMERICAN FOUNDATION
FOR THE BLIND INC.

INTELLIGENCE TESTING

METHODS AND RESULTS

BY

RUDOLPH PINTNER

PROFESSOR OF EDUCATION IN TEACHERS COLLEGE
COLUMBIA UNIVERSITY

NEW EDITION

NEW YORK
HENRY HOLT AND COMPANY

Translated into Greek
by
Const. Myrianth, 1927

PRINTED IN THE
UNITED STATES OF AMERICA

PREFACE TO THE SECOND EDITION

Since the appearance of the first edition of this book in 1923, the interest in intelligence testing has increased and widened to such an extent as to make a revision of the book necessary. New and better tests have appeared, much has been written about the theory of intelligence, and the tests themselves have been given to greater numbers and more varied groups of individuals. This has meant the complete re-writing of many chapters and the inclusion of entirely new chapters. The bibliographies at the end of each chapter have been brought up to date and lengthened, so as to provide a good starting point for a more intensive study of the topics discussed.

R. P.

Yonkers, New York,
October 16, 1930.

PREFACE TO FIRST EDITION

This book is an attempt to give a simple account of intelligence testing and the results which have so far been achieved by the testing movement. It is designed for use as a text in a college course, and it is hoped that it will prove useful in serving as a guide to the thousands of teachers who are now becoming interested in the use of intelligence tests in their schools.

The book is not a treatise on measurement in education or psychology. For this purpose, we have already the valuable works of Thorndike, Rugg, and others. Nor does it deal with the technique of test construction, which has been recently covered by McCall. Furthermore, it does not deal with educational tests, that is, with tests of achievement in school subjects, for in this field there are now many books. It is rather an attempt to tell the reader what is meant by intelligence testing, what means are employed to test general intelligence, and what results have been achieved.

Part One is mainly historical and theoretical in nature. It shows the gradual evolution of the intelligence test and discusses some of the basic assumptions underlying the work.

Part Two is a description of the various tests, individual and group, that are at present available. It aims to give the student a survey of the various methods by means of which intelligence is tested. The classroom instructor will supplement here by giving or demonstrating to his class the different types of tests. In some classes practical work with the tests in schools will be feasible and desirable.

Part Three summarizes the main results of intelligence testing. Up to the present time this material has been scattered in numerous periodicals, monographs and books. Many of these are inaccessible to the student. The author has attempted to bring this material together and indicate what conclusions can be drawn at the present time. The numerous references in this section may be utilized by the instructor as assignments for further study by his students.

Bibliographical references follow each chapter. These are arranged in alphabetical order according to the name of the writer. In the text, the date of publication of the work referred to is put in parenthesis after the writer's name.

Throughout the book the influence of Thorndike's work will be obvious to the student. No one man has had more to do with stimulating the measurement movement in this country than Professor Thorndike. It is fitting, therefore, that the author should acknowledge his indebtedness in this place, both for the inspiration received from Professor Thorndike's writings and for his personal advice and encouragement. He also wishes to acknowledge the very valuable help rendered by Professor Peter Sandiford of Toronto University, who read the original manuscript and made many important suggestions. Also, he takes pleasure in acknowledging here the great assistance rendered by his wife in preparing the manuscript for the printer and in seeing the book through the press.

Intelligence testing is of recent growth, and the accomplishments of the last fifteen years have been great. Nevertheless, we stand merely at the beginning of an important chapter in the measurement of human behavior. The future will see new and more accurate tests. It will see tests applied for purposes now unthought of. The types of individuals tested, which form the chapter headings of Part Three, will undoubtedly increase in numbers. A future edition of such a book as this may well have such chapter headings as "The Immigrant"; "The Voter"; "The Applicant for a Marriage License"; "The Candidate for Public Office"; "The Civil Servant"; and so forth. For intelligence is one of the aspects of an individual's personality that is of great importance in modern civilized life. And now that we are able to measure it with a fair degree of accuracy, intelligence tests will find a wider and wider field of application.

R. P.

New York,
January, 1923.

CONTENTS

PART I. HISTORICAL AND THEORETICAL

PART II. THE METHODS

PART III. THE RESULTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

PART I

HISTORICAL AND THEORETICAL

CHAPTER I

EARLY HISTORY OF INTELLIGENCE TESTING

Although the intelligence test, as we know it to-day, is of recent growth, it is, nevertheless, interesting to trace its early history, and to find some of the causes which led to its development. Like so many of the results of modern science, the intelligence test may be said to have appeared as the fulfilment of a need that existed. It came to supply a want in society. And the science of psychology had progressed far enough in the problem of mental testing to be ready to fill this need when the time came. We may thus think of these two aspects, the theoretical interest and the practical need, and we may further consider them as brought to a focus in the work of Binet, as exemplified by his construction of the Binet-Simon Scale. Let us call the theoretical interest in the problem the psychological interest and the practical interest the sociological interest. The following diagram will help to illustrate this presentation of the subject:

A. SOCIOLOGICAL FACTORS Insane Deaf Feebleminded Delinquents		*X. APPLICATIONS* Feebleminded Delinquents School children College students
	Binet (1857-1911)	
B. PSYCHOLOGICAL FACTORS Experimental Psychology Individual Differences Eugenics Anthropology		*Y. TESTS* Binet Revisions Stanford scale Performance scale Group tests

The left side of our diagram represents the forces, practical and theoretical, leading up to the work of Binet, and the right side the modern development of the work, which we have represented as brought to a focus by Binet. The upper half of the diagram represents the practical field, and the lower the theoretical. The sociological trend stimulated the psychologist to devise means to fill the needs of society and once these means were supplied, we see the application of these to many different groups of individuals. The psychological trend gave the psychologist the tools with which to work and once these were adapted to the new demands, we see an extension of these kinds of tools in the shape of innumerable scales and tests, only a few of which we have indicated in our diagram.

By bringing these trends to a focus in the work of Binet, we do not mean to suggest that he was the only worker in the field, nor that his influence was the only influence leading to the development of scales for mental measurement. There were many workers and many currents of interest, so many, indeed, that our diagram would become too complicated if we tried to include them all. It was, however, the work of Binet during the years 1905 to 1911, around which the main interest centered, and it was his scale that first won universal recognition as a practical means for the measurement of mental ability.

In this chapter we shall trace briefly the two influences, the sociological and the psychological, that led to the establishment of scales for the measurement of intelligence. The work of Binet and of the other workers who followed him, and the results of the application of their tests will be taken up in succeeding chapters.

A. THE SOCIOLOGICAL TREND

By this we mean those forces in society directing our attention to the defective and delinquent classes. These classes first aroused the interest of the psychologist because of their peculiarities in mental make-up. This is particularly true of the

feebleminded, and the application of the first scales for measuring intelligence was largely restricted to them. It might be well, therefore, to trace briefly the attitude of society towards this group of individuals.

Ancient Period.—We have little information as to the feebleminded in ancient Greece and Rome. Exposure of undesirable children was practised and at certain periods and in certain places was common. In general the obviously physically defective were exposed, and, to the extent that feeblemindedness is accompanied by physical defect, the feebleminded would thus be eliminated. The great mass of feebleminded, sound in body, would escape. During the period of the Roman Empire exposure was not so common, and the cynic has suggested that, had the old Roman custom more commonly prevailed, the world might have been spared the excesses of a Nero or a Commodus.

Medieval Period.—The emphasis of Christianity upon charity and mercy caused a decided change in the attitude of society towards the physically and mentally defective. All classes of the "despised and rejected" were cared for in asylums by the Church. This attitude was marked by sympathy and pity, but was absolutely lacking in understanding or in helpful service. The insane and feebleminded were tolerated and in many cases regarded with a sort of religious awe. They were supposed to be under the special protection of God, and the ravings of the "fool" were sometimes taken for divine revelations. This religious reverence is well expressed in the French "enfants du bon Dieu," and in the term "innocents" as used in Scotland and Ireland. Later on the "fool" graduated into the "jester," and this position was often filled by men of high intellectual attainments.

During this period the laws of England differentiate between the idiot (*idiota a nativitate*) and the lunatic (*idiota a causa et infirmitate*). In an old law book, "The New Natura Brevium," by Sir Anthony Fitzherbert, published in 1534, we find the following suggestion of primitive intelligence tests: "And

he who shall be said to be a sot and idiot from his birth, is such a person who cannot account or number twenty pence, nor can tell who was his father or mother, nor how old he is, etc., so as it may appear that he hath no understanding of reason what shall be for his profit, nor what for his loss. But if he hath such understanding, that he know and understand his letters, and do read by teaching or information of another man, then it seemeth he is not a sot nor a natural idiot."

The Renaissance.—The emphasis of Protestantism upon the individual and upon his individual responsibility for his deeds and misdeeds, caused a sudden change in the attitude of society towards the feebleminded and insane. Far from being considered the special protégés of the Deity, they are now regarded as the children of Satan. They are possessed of devils and strenuous treatment must, therefore, be applied to drive out the devils. What they do, they do intentionally and of their own free will, and, therefore, they must be chastised for their misdeeds. So the insane and feebleminded who offend are whipped and bound in chains and cast into dungeons until such time as they reform. This period has been well called "the era of whips and chains."

The industrial revolution of the eighteenth century produced a great demand for child labor. The condition of the child mill-worker in all industrial centers was pitiable. An interesting reference to the feebleminded appears in the custom, prevalent in England in the eighteenth century, of binding out pauper children to the mill owners. "The parish authorities, in order to get rid of imbeciles, often bargained that the mill-owners take one idiot with every twenty children. What became of the idiots is not known, but in most cases they did not last long and mysteriously disappeared" (Payne, 16).

Modern Period.—Gradually the modern attitude developed, characterized by a scientific interest in the insane and feebleminded, and also by a deeper sense of social justice to all classes. The fight for the better treatment of the insane was prosecuted vigorously during the nineteenth century. It di-

rected the attention of physicians and psychologists to this class and one of the results was the beginning of a real study of insanity. As this went on, the interest was extended to the feebleminded and the present era of enlightened care and definite study of feeblemindedness was begun.

The Deaf.—Excluding from consideration the charitable asylums that were found in early times and in the middle ages, caring for the outcasts and the defectives in general, we may say that, of children who are not entirely normal, deaf children were the first to attract the special interest of the educator. The affliction of deafness is more striking and more dramatic than the afflictions of blindness or of feeblemindedness. The uneducated deaf and, therefore, dumb child is so utterly unlike the normal hearing child as to arrest attention at once. He has no means of communication with his fellows and cannot say the simplest words, a condition that is only approached by a very low grade idiot, whereas in other ways he shows signs of intelligence much above the idiot. Again the crude attempts of the deaf to communicate with others by means of simple natural signs must have suggested to the educator a simple means of instruction that needed only to be enlarged and systematized.

Ponce de León (1520-1584) is credited with being one of the first to educate the deaf by signs and also by oral speech. In the seventeenth century Juan Pablo Bonet is supposed to have invented a manual alphabet from which that now used in the United States is descended. In the eighteenth century Pereire in France revived Bonet's alphabet and taught lip-reading and speech. The Institution for Deaf Mutes in Paris was founded by the Abbé de l'Epée (1712-1789), and the methods employed in this school were later on studied by T. H. Gallaudet and brought over to America. He became the head of the first school for the deaf in this country, the American Asylum for the Deaf and Dumb at Hartford, Connecticut, founded in 1817.

The Blind.—After the work for the deaf was well started it was natural that the interest of educators should extend to the blind. Haüy in France began definite constructive work in

1798. Very soon this line of work extended to several countries and Dr. Howe in Boston was the first superintendent of the Perkins Institute for the Blind founded in 1833.

Interest being thus directed toward abnormal children, it could not be long before the mentally defective child, as contrasted with those suffering primarily from the physical defects of blindness and deafness, would claim attention. It is a fairly safe conjecture to suppose that in the actual work with the deaf and the blind the problem of mental defect would arise, and we know that idiots and imbeciles were sometimes admitted to deaf and dumb asylums. The interest of the educators, however, was rightly centered on the handicaps of deafness and blindness, and it was, therefore, left for a more dramatic incident to turn the attention of scientists to the problem of feeblemindedness itself.

The Feebleminded.—In 1797, a so-called wild boy was found by some hunters in the woods of Caune, in the Department of Aveyron in France, and was brought to Paris, where he aroused considerable interest. He is known in our literature as the Wild Boy of Aveyron (*Juvenis Averionensis*), and we may consider him the first feebleminded child whose education was scientifically attempted. That he was an imbecile was not admitted by many of the scientific men who studied him, although Pinel suspected this to be the case. Pinel was Physician in Chief to the Bicêtre in Paris, a charitable institution or foundling hospital, founded by St. Vincent de Paul, caring for the feeble in mind and body. The boy, however, aroused the interest of the followers of de Condillac, the sensualist philosopher, and they imagined that here would be an admirable opportunity to watch the effect of sensations upon ideas in the gradual transition from savagery to civilization. There was no one more fitted to undertake this task than Itard, the philosophically minded physician of the Institution for Deaf-Mutes. Itard was familiar with what had been accomplished by Pereire and de l'Epée, and was daily in touch with the work of Sicard at the deaf school. He was well qualified, therefore,

to make good use of the methods of education that were being used with the deaf.

Itard worked hard and accomplished much but could not restore the boy to complete normality. When it dawned upon Itard at last that the boy was an imbecile or idiot, he threw up his work in despair. He shared the common medical belief that idiocy was incurable and that idiots were "human brutes" separated by a great gulf from the normal, and, therefore, once sure of a diagnosis of idiocy there was nothing to do but to send the boy to the Bicêtre as a custodial case unworthy of training. Itard failed to see that, although the boy was an idiot or imbecile, his course of training had been of great good to him and had made him less of a burden to society than he had been before.

Seguin.—The great step in advance made by the method of education followed by Itard and the great fact of the improvability of the feebleminded was seen and appreciated by Seguin, the pupil of Itard. He had watched with interest the experiment of his master and he carried on the work from the point where Itard abandoned it. In 1837 Seguin commenced the training of a few cases of feebleminded children. In 1842 he convinced the authorities of the desirability of educating the idiots and imbeciles at the Bicêtre and he was put in charge of the school there. This marks the beginning of state schools for the feebleminded. From this time on the State recognizes the necessity for training these individuals in addition to merely housing and feeding them.

Seguin is also important in our history because he is the author of the first standard book dealing with the education and treatment of the feebleminded. His book, "Traitement moral, hygiène et éducation des idiots," was published in 1846, four years after he commenced his work at the Bicêtre. This book has been well called "the emancipation proclamation for the fettered soul of the idiot" (Johnson, 95). There is much in it that we do not agree with at the present time, but on the other hand it shows a marvelous insight into many of the

aspects of our subject. We do not agree with the sharp distinction that he draws between idiocy and imbecility and between both of these and backwardness, but his plea for the regulation of marriage to restrict the propagation of the feebleminded shows that he appreciated as well as we do in the twentieth century the importance of the factor of heredity. The whole book is written in an aggressive style and many passages are directed against the ignorance, apathy and conservatism of the medical profession. He feels keenly that the medical profession as a whole is not as actively interested in the problem of the feebleminded as it should be, and that their conservative attitude and their pretence of knowing all about it stands in the way of progress. Seguin's whole life was devoted to work with the feebleminded and by coming to America in 1848 he stimulated directly the work in the United States. While in this country he published in 1864 his English book, "Idiocy; Its Diagnosis and Treatment by the Physiological Method."

In other countries as well as in France, attention was being directed toward the feebleminded, and England, Germany and Switzerland soon took up the work. The first state institution for the feebleminded in America was opened in Massachusetts in 1849. New York followed in 1851, and from this time onward up to the present time we have a record of the opening of one institution after another. The great majority of countries now fully recognize the duty of making some kind of provision for the mentally defective.

Special Classes.—Commencing a little later than the establishment of institutions and much slower in making headway, was the movement for the separation of the backward and dull children into special classes in the public schools. The first special class was started in Halle, Germany, in 1859, with the idea of stimulating the child in order to put him back into the regular classes. This idea existed for some time, but finally had to be abandoned as the classes gradually filled up with feebleminded children who could not keep pace, let alone catch up with the regular classes. In the United States, the first

special class for backward children seems to have been organized in Cleveland in 1893 (Mitchell, 16). Although the first classes were started only shortly after the opening of the first institutions, it was not until the first decade of the twentieth century that the special class became a common feature of the ordinary public school system.

Psychological Clinic.—We thus see a growing interest during the nineteenth century on the part of the social reformer and the educator in the care of the feebleminded and it is towards the end of the century that we notice the first definite interest on the part of the psychologist. This is typified by the opening of the Psychological Clinic in the University of Pennsylvania in 1896 under Dr. Lightner Witmer. Here for the first time we see the emphasis placed upon the necessity of a careful psychological diagnosis of the nature of the mental deficiency together with an attempt to treat such deficiency as far as it may be amenable to treatment (Witmer, 07). Let us, however, retrace our steps for the time being in order to follow other influences in psychology which were leading up to the development of clinical psychology and the measurement of mentality.

B. THE PSYCHOLOGICAL TREND

If we trace the influence in psychology and allied sciences that led to an interest in mental testing, we can for convenience distinguish four lines of approach: (1) experimental psychology; (2) the study of individual differences; (3) the growth of eugenics and (4) anthropological measurement.

1. *Experimental Psychology.*—The growth of experimental psychology from the opening of Wundt's laboratory in Leipzig in 1879 was very rapid. The apparatus and methods of physiology and physics were pressed into the service of the new science. The main interest centered upon a study of the general laws of the normal human mind, and, in arriving at these, the differences between the different individual observers acted as disturbing factors. It was not long, however, before these

differences became in themselves of interest. This is, perhaps, best seen in the study of reaction to a visual, an auditory, a cutaneous or some other kind of stimulus regardless of the observer's own peculiarities. The differences between observers, however, soon led to the study of different types of reaction, and we have then the many studies dealing with so-called sensory, motor and mixed types. It can readily be seen, therefore, how at an early date a study of individual differences themselves was bound to arise.

Before proceeding to touch on this factor, we may note an outgrowth of German experimental psychology that indirectly helped the growth of clinical psychology. This was the carrying over of the apparatus and methods of the psychological laboratory into the psychiatrist's examining room. The German psychiatrists were more influenced than those of any other country by the new experimental psychology, and many of them received a thorough training in the psychological laboratory. Kraeplin, Sommer and Ziehen are the outstanding names and the work they did was essentially the application of the methods of experimental psychology to the study of insanity. Although this work contributed interesting facts to our knowledge of the psychology of insanity, it failed to give us any new viewpoint, such as the Freudian conception of the abnormal mind or Binet's concept of general intelligence and the possibility of its measurement.

2. *Individual Differences.*—We have seen how a study of individual differences arose out of the work in the psychological laboratory, and in the last decade of the nineteenth century a great many studies appeared. In this country the work was championed by Cattell and it would seem that we are indebted to him for fixing the word "test," as denoting a simple task to be performed by subjects in the investigation of individual differences. Already in 1890 we find him writing on "Mental Tests and Measurements," pleading for standardization of methods of procedure, and urging the necessity for the establishment of norms.

Very few of the tests described by Cattell are tests of general intelligence; they are mainly sensory and sensory-motor tests. Nevertheless, we see here a beginning of the type of work that led later to intelligence testing. In 1894 Cattell began testing the students of Columbia College during their first and fourth academic years, and thus began the testing of Columbia students, which has culminated in the work of Thorndike at the present time. In 1896 Cattell and Farrand published the first results of these tests. About one hundred students were given tests of vital capacity, strength of grip, vision, reaction time, pain, memory, imagery, etc. The authors urge the adoption of these tests and that they be given in a standard way by all psychological laboratories.

In this same decade we have other articles showing how the interest was growing and how the idea was expanding. Bolton (92) gives memory tests for digits to a great many children, studying the growth of memory and comparing the results in memory with the intellectual acuteness of the children. To be sure the "intellectual acuteness" is merely based upon the judgment of the teacher, but it is already becoming a factor in the study of individual differences.

In the Columbian Exposition in Chicago in 1893 Jastrow gave tests to visitors who were interested. He had a great number of tests, sensory, sensory-motor, memory and reaction tests. Undoubtedly a great many people saw these tests, and so this exhibit helped along the idea of mental testing. During the same year Gilbert gave a battery of eight tests to 1,200 children. He gave age norms for these tests and showed how the scores increased up to puberty. Furthermore, he compared the test results with teachers' estimates of the pupils' "general ability." He found a slight relationship between some of the tests and these estimates.

Again it is interesting to note an article by Bourdon (95) because it contributed a test that has since been widely used. Bourdon's interest was philosophical rather than practical. He was studying the rate and time of perception and in so doing

he used the cancellation of one letter on a printed page as a test for his observers. He calls it a test of discrimination. This seems to be the first use of a cancellation test, a type of test that has been used extensively and one that lends itself to infinite variations.

Johnson (95) shows how the influence of the Child Study Movement, stimulated by Stanley Hall, was spreading over to an interest in feebleminded children. His work is historical, pedagogical, and observational in character, but also includes several psychological tests. The performances of feebleminded and normal children on the tests are compared. He has no adequate norms for ordinary children, but his work was pointing the way to the necessity for standards.

In 1897 we have the first appearance of the now famous Ebbinghaus completion test. Ebbinghaus was working on the problem of mental fatigue in school children, and so he devised three tests to be applied for a few minutes during each period of the school day. These tests were: (1) rapid calculation; (2) memory for digits; (3) completion of sentences. The completion test proved very successful. Ebbinghaus thought that it possessed the essentials of a good test of intelligence. Intelligence for him consisted of "combination ability," the ability to combine parts into a whole, synthesis rather than analysis. The completion test has indeed proved very valuable and it is in use to-day in a great many intelligence tests.

In Sharp's (98) work we see the direct influence of Binet and Henri in the application of certain tests in memory, imagery, imagination, attention, etc., to about eight students. We shall reserve for a later section a description of the work of Binet, but it is interesting to note at this early date the influence of his work, and this we must remember is long before the publication of the Binet-Simon Scale.

By 1900 we note the interest extending to the testing of school children. Kirkpatrick tests 500 children with tests of counting aloud, making vertical marks, sorting cards, interpreting ink spots and so on. The results of these tests are com-

pared with the grades given the children in school work. There is no elaborate attempt on the part of the author to be concerned with just what mythical faculty of the mind each one of his tests seems to be testing. He is anxious to see in a general way whether these tests are indicative of good accomplishment in school work. He feels greatly the need of standards and makes a beginning of standardization by age.

The work of testing students under Cattell's direction was going on all this time at Columbia University and in 1901 we have a report by Wissler (01) of some of the results. In this report the Pearson correlation technique seems to have been used for the first time for the comparison of test with test, and tests with college grades. Naturally all the correlation coefficients were extremely low, because each test was considered separately, because the tests were very short and largely sensory-motor in character and because the college students were a highly selected group. This report dampened the ardor of the psychologists interested in tests. It seemed as if there were no relation between mental functions and general ability. The idea of pooling tests had not yet arisen.

In 1903 appeared one of the most interesting articles prior to the Binet Scale. This is Kelly's article on "Psycho-physical Tests of Normal and Abnormal Children." It is interesting chiefly because he clearly states as his purpose the attempt to find a simple method of differentiating between normal and abnormal children. His tests are largely physical, and he feels keenly the need of norms. He finds in general an increase in motor co-ordination as intelligence increases; also the lower the intelligence, the more prominent the element of fatigue.

The most significant work in this country, before the appearance of the Binet Scale, is undoubtedly the tests of feebleminded children conducted by Norsworthy (06). At this early period we find her giving what are, in essence, group tests of intelligence and expressing the standing of the child in terms of the variability of the group. This is really what many group tests are doing at the present time.

We see, therefore, from this survey of a few of the studies in individual differences that were appearing at the beginning of the twentieth century, that interest was being aroused in the possibility of the measurement of intelligence and that the time was ripe for the work of Binet.

3. *The Eugenics Movement.*—Another great stimulus to the study and measurement of individual differences came from the Eugenics Movement. This movement was started and fostered by Sir Francis Galton, one of the most versatile of British scientists in the latter half of the nineteenth century. His book, "Hereditary Genius," published in 1869, begins with the following significant sentence: "I propose to show in this book that a man's natural abilities are derived by inheritance, under exactly the same limitations as are the form and physical features of the whole organic world." This emphasis upon "natural abilities" or mental traits is for us the important thing, because attention could not long be focussed upon the inheritance of such abilities, without the necessity soon arising of some method of evaluating or measuring the amount of such abilities that may be inherited. Indeed, in this very first book by Galton we find him constructing an imaginary scale for the measurement of general ability. This imaginary scale is based upon the theory of a normal distribution and ranges from the lowest idiot to the highest genius. Here Galton introduced a valuable quantitative concept, and began the breaking-down of the commonly accepted idea of the existence of specific types, such as idiots and geniuses. People differ from each other in general ability by measurable amounts and cannot be grouped into several distinct and specific types.

Galton elaborated this idea of a scale and divided it into fourteen grades. He assumed without objective tests that races varied in intelligence, and he suggested that negroes are very probably two grades below whites. The highest race or group he considered the Athenians of classical times and he placed them two grades above modern Europeans.

In all Galton's work we note this interest in the measurement

of ability, although he himself did not devise any specific tests for it. His idea of general intelligence has certainly had influence upon psychological thought and it is in many respects very much like the ordinary psychological conception at the present time. For instance, in his "English Men of Science" (74), in discussing the presence or absence of what he calls the innate tendency for science, he says, "Nay, further, it appears that of the men who have no natural taste for science and yet succeed in it, many belong to gifted families, and may therefore *be accredited with sufficient general abilities* to leave their mark on whatever subject it becomes their business to undertake."

In 1884 Galton founded his Anthropometric Laboratory for many different types of measurement and it is interesting to note that reaction time is found among his list of tests. In the mathematical handling of the data he expanded the idea of percentile grades, suggested to him by the work of Quetelet. Later on, in 1886, he introduced and applied the coefficient of correlation in a simple form. About 1901 we note the founding of the Biometric Laboratory by Karl Pearson in University College, London, and in 1905 the founding of the Eugenics Laboratory by Galton.

These events mark the firm establishment of the concept of mental measurement, and the work of the British Biometric School has contributed greatly to the mathematical handling of data and particularly to our understanding of the theory of correlation.

4. *Anthropological Measurement.*—Our fourth and last line of approach we shall deal with very briefly. The field of anthropology is, of course, very vast and the number of different kinds of measurements made by anthropologists very great. Only the measurements of the skull, however, interest us here. It was inevitable, when such measurements were made, that interest should be aroused as to the abilities of the individuals measured. Head measurements of all kinds were made and various cephalic indices calculated. When the ques-

tion as to the relation of such to the ability of the individuals measured was raised, it led to attempts to get quantitative statements of ability. In this way considerable stimulus was given to devising and giving psychological tests. Binet himself at first thought that a short cut to the evaluation of an individual's mental ability might be obtained by some appropriate cephalic index. We shall see later on that he abandoned this idea.

In addition to the anthropological interest in the measurement of the head, we must consider the study of the head made by the much-discredited science of phrenology. It started from a perfectly legitimate and disinterested study of the localization of the functions of the cerebrum, continued with extravagant theorizings unsupported by experimental evidence, and ultimately fell into the hands of quacks and charlatans. The existence of this pseudo-science and the belief in it at the present time is surprising. We mention this merely as an indication of the perennial human interest in the abilities and capacities of the human race and the all-absorbing desire of knowing something more definite about them. This interest is common to the psychologist who is constructing elaborate mental tests for prognostic purposes, as it is to the country bumpkin who pays a quarter to have his head examined to see what he is fit for and destined to become.

Conclusion.—We have attempted in this chapter to follow the two lines shown on the left half of our diagram on page 3, indicating the practical and theoretical forces that were at work leading to interest in and experimentation with mental tests, and culminating in the work of Binet in the first decade of the twentieth century. We have called the first the sociological trend because it deals with the attitude of society towards those of abnormal mental characteristics. In this sketch we have largely concentrated upon the feebleminded, because it was the feebleminded who first presented the problem of mental measurement as a definite task to the psychologist. The second, the theoretical or psychological approach,

has shown us how the interest of psychology in the behavior of man naturally led to a study of individual differences and to an attempt to measure these differences, culminating with the mental test as we know it to-day.

And back of all this is the perennial interest of mankind in its own capacities, an interest that shows itself in phrenology and fortune-telling, in palmistry and graphology as well as in the laboratory or the clinic of the psychologist.

REFERENCES

Bolton, T. L. (92). "The Growth of Memory in School Children." *Am. J. of Psych.*, Vol. IV, pp. 362-380.

Bourdon, B. (95). "Observations Comparatives sur la Reconnaissance, la Discrimination et l'Association." *Revue Philosophique*, Vol. XL, pp. 153-185.

Cattell, J. McK. (90). "Mental Tests and Measurements." *Mind*, Vol. XV, pp. 373-380.

Cattell, J. McK. and Farrand, L. (96). "Physical and Mental Measurements of the Students of Columbia University." *Psych. Review*, Vol. III, pp. 618-648.

Ebbinghaus, H. (97). "Uber eine neue Methode zur prüfung geistiger Fahigkeiten und ihre Andwendung bei Schulkindern." *Zeitschr. f. angewandte Psychol.*, Vol. 13, pp. 401-459.

Fernald, W. E. (93). *The History of the Treatment of the Feebleminded.* Twentieth National Conference of Charities and Correction.

Fitzherbert, Sir Anthony (1793). *The New Natura Brevium.* Dublin. (First Ed. 1534.)

Galton, F. (69). *Hereditary Genius.* New York, 1880 (Reprint).

Galton, F. (74). *English Men of Science; Their Nature and Nurture.* New York, 1890 (Reprint).

Gilbert, J. A. (94). "Researches on the Mental and Physical Development of Young Children." *Stud. of Yale Psychol. Lab.*, Vol. 2, pp. 40-100.

Johnson, G. E. (95). "Contribution to the Psychology and Pedagogy of Feebleminded Children." *Ped. Sem.*, Vol. III, No. 2, pp. 246-301.

Kelly, R. L. (03). "Studies from the Psychological Laboratory of the University of Chicago. Psycho-physical Tests of Normal and Abnormal Children." *Psych. Review*, Vol. X, pp. 345-372.

Kirkpatrick, E. A. (00). "Individual Tests of School Children." *Psych. Review*, Vol. VII, pp. 274-280.

Mitchell, D. (16). *Schools and Classes for Exceptional Children.* Cleveland Education Survey.

Norsworthy, N. (06). "The Psychology of Mentally Deficient Children." *Archives of Psychology,* No. 1.

Payne, G. H. (16). *The Child in Human Progress.* New York.

Pearson, K. (24). *The Life, Letters and Labours of Francis Galton.* Two volumes, Cambridge University Press.

Peterson, J. (25). *Early Conceptions and Tests of Intelligence.* Yonkers.

Seguin, E. (46). *Traitement Moral, Hygiène et Education des Idiots.* Reprinted 1906. Paris, Alcan.

Seguin, E. (64). *Idiocy: Its Diagnosis and Treatment by the Physiological Method.* Revised 1866 under title *Idiocy and Its Treatment by the Physiological Method.* Reprinted by Teachers College, Columbia University, 1907.

Sharp, S. E. (98-9). "Individual Psychology. A study in Psychological Method." *Am. J. of Psych.,* Vol. X, pp. 329-391.

Wissler, C. (01). "The Correlation of Mental and Physical Tests." *Psych. Monographs,* Vol. III, No. 6, Whole No. 16.

Witmer, L. (07). "Clinical Psychology." *Psych. Clinic,* Vol. I, No. 1, pp. 1-9.

Young, K. (24). "The History of Mental Testing." *Ped. Sem.,* Vol. 31, pp. 1-48.

CHAPTER II

THE WORK OF BINET

The work of Binet is important and merits special consideration because of the great stimulus he gave intelligence testing by the construction of the famous Binet-Simon Scale for the measurement of intelligence. It is of great importance historically because so much of the later development of intelligence testing is implicit in his work. We are still elaborating upon the ideas that he set forth, and his concept of intelligence is essentially the one that is held at the present time by psychologists. This does not mean that measurement of intelligence would not have been attempted without the work of Binet. It certainly would, and we have tried to show in the previous chapter the forces that were acting in that direction. The work in mental tests started by Cattell, and particularly the work of Thorndike in educational measurement, would undoubtedly have culminated in the testing of intelligence as we know it to-day. In England the work of Burt was also distinctly tending in the same direction. But it seems unquestionable to the writer, that without Binet this development would have been much slower and would probably not have taken the decidedly practical turn at the outset which the work of Binet gave it.

Biography.—Alfred Binet was born in Nice, France, on July 11th, 1857. His father was a physician and his mother an artist. He was a pupil at the Lycée Saint Louis, Paris. He later received his *licencié en droit* in 1878. He then studied medicine, and worked particularly under Charcot and Féré. To the influence of the former he undoubtedly owed his knowledge of and interest in abnormal psychology. During this

period he published a memoir, *"Sur la vie psychique des micro-organismes,"* and also "The Psychology of Reasoning." In conjunction with Féré he published "Animal Magnetism" and later on, in 1891, he wrote "Alterations of Personality." In 1889 Beaunis and Binet founded the Psychological Laboratory at the Sorbonne. In 1890 Binet received his *licencié ès sciences naturelles,* and in 1892 he became adjunct director along with Professor Beaunis of the Laboratory of Physiological Psychology at the Sorbonne. Later on, he himself became Director of the Laboratory and held this position until his death. In 1894 he received his *doctorat ès sciences* with a thesis entitled, *"Contribution à l'étude du système nerveux sous-intestinal des insectes."* The next year, in 1895, he founded *"L'année psychologique,"* which was to prove the main avenue for the publication of his future work. Indeed, many of the thick volumes of this periodical are largely made up of the writings of Binet himself. One is forced to marvel at his capacity for work. A perusal of the volumes of the *Année* is the best means of tracing the growth of his interest in intelligence testing, culminating in 1905 with the first set of tests, and followed thereafter by the elaboration of the Scale in 1908 and the revision of the Scale in 1911, the year of his death. In 1900 the *"Société libre pour l'étude psychologique de l'enfant"* was founded, an association of psychologists and school teachers who worked on practical problems in the schools under the leadership of Binet. In 1904 the Minister of Public Instruction appointed Binet a member of a commission made up of medical men, educators and scientists to formulate recommendations for the administration of special classes in the public schools. As a member of this commission Binet did excellent work in spite of the opposition of many of the physicians. His work on the commission stimulated him to put his tests to practical use. Along with Simon he tried them out in the school and thus appeared the first rough scale in 1905.

Binet's death at the age of 54 on October 11th, 1911, leaves one with the feeling of a great loss to the science of psychology

because of the conviction that he would have continued and developed the concepts inherent in his work. We have unquestionably lost the extension by him of these concepts to the field of psychiatry, a line of thought that he was following out at the time of his death.

Measurement.—It may repay us, therefore, to make a brief survey of some of his articles as they appeared in the *Année*, dwelling more particularly on those that are directly or indirectly connected with the idea of mental measurement. As early as 1896 we find him collaborating with Henri in an article discussing the field of individual psychology. In 1898 he contributes an article to the *Revue Philosophique* with the significant title, "Measurement in Individual Psychology." The question of measurement, he says, is the important thing—"How can we measure richness of inspiration, accuracy of judgment and the general ability of the mind?" Present-day psychology is gradually answering this question that Binet propounded more than thirty years ago. It is interesting, moreover, to note the tests that he mentions in this same article. Some of these are:—drawing a square from memory; suggestibility to length of lines; memory for numbers; rearrangement of dissected sentences; answers to questions involving moral judgments; comprehension of an abstract passage; folding paper test. We see here the beginnings of the tests that later on proved so useful in his scale. It is not measurement in the physical sense, says Binet, but classification of individuals with reference to others. This again is one of the fundamental concepts of mental measurement. Indeed this article of Binet embodies surprisingly the gist of present-day thought on mental measurement, and looking back from our present position it would seem as if nothing could have been more simple than for Binet to march straight on toward the scale of tests. As a matter of fact, however, there were many difficulties to be overcome and the way that he actually took was long and devious. Before he adopted the simple

tests foreshadowed here, he tried many other possible means of measurement.

The next year he published an elaborate study on the consumption of food in relation to mental work, and in the same year also we note an article on anthropometrical measurements of abnormal boys by Simon, a man who later became so closely identified with Binet. This was a comparison of measurements of the waist, thorax, head, etc., with those of normal boys. It may be interesting to notice that in this article Simon refers to his doctor's thesis entitled, "Documents pertaining to the correlation between physical and mental development" (1900), showing that Simon early became interested in the problem of mental development.

Attention and Adaptation.—In this same volume also we have Binet's long contribution on *"Attention et Adaptation,"* his most important work prior to the scale proper. His attempt here, he says, is to measure voluntary attention or more particularly to study voluntary attention in relation to intelligence and as a means for distinguishing differences in intelligence. For Binet, to measure is by this time synonymous with giving tests, and we, therefore, see him once again exercising his ingenuity in the devising of simple tests. This time, among others, we note designs from memory, and the truncated pyramid and Greek key pattern, which now have become famous in the Binet Scale and its adaptations, appear here. He also uses dot counting, simultaneous adding, cancellation, reaction time and the like. These tests he gives to two groups of children of known intelligence. One group is bright and the other dull, or intelligent and unintelligent, as he calls them. The intelligent, he says, show a quicker and a better adaptation. "We have implicitly admitted that attention consists in a mental adaptation to a situation which is new to us." How surprisingly close to Stern's well-known later definition of general intelligence does this thought of Binet come.

Head Measurements.—The next year we see Binet going off on one of his numerous tangents into a lengthy piece of work

on the technique of the measurement of the living head. Our chief interest in this work lies in the fact that he compares the head measurements of intelligent and unintelligent pupils. The differences between the averages of such groups are always in favor of the intelligent group, but then the differences are small and the overlapping great. The more carefully the groups are selected the greater the differences become. The intelligent child is distinctly superior. Nevertheless for individual diagnosis such measurements are useless.

In spite of this negative result so far as our main interest is concerned, the next year of Binet's publications shows nothing but a number of articles all dealing with the results of his cephalometric research. He plunges into the question of the growth of the head in children between the ages of four to eighteen. He compares the head measurements of the blind and the deaf with those of normal seeing and hearing individuals.

Seeming to have exhausted his interest in this direction, we see him in the following year comparing intelligent and unintelligent groups of children with reference to their sensibility as measured by the two-point threshold on the skin. The intelligent have a larger percentage of correct responses. He believes these experiments with the aesthesiometer are a good measure of voluntary attention.

Higher Mental Processes.—In 1903 appeared his book "*L'étude expérimentale de l'intelligence.*" Here he takes intelligence in a broad sense as equivalent to the higher mental processes. He does not theorize much but details all his attempts at measurement with innumerable tests. In the main he uses as his two observers his two daughters, thirteen and fourteen years old. Among the tests used we find the completion of sentences, the description of an object and the description of a picture. The picture used is that of a peddler and his boy pulling a wagon, the same picture used later in his scale and revived in the Yerkes-Bridges Scale.

Handwriting.—In 1904 appear a few minor articles showing the wide range of his interests. One deals with a psychological analysis of the writings and mental make-up of a French writer, Paul Hervieu. The other is a study of graphology and raises the question as to whether handwriting experts can determine sex, age and intelligence. Note again how almost all his work brings in the question of intelligence in some form or other. He is rather favorably disposed towards the position that handwriting is a crude index of intelligence.

Feeblemindedness.—Our brief survey up to this point of the articles published by Binet has made it sufficiently clear that the measurement of intelligence had been one of his main thoughts at least for the past ten years. Almost every piece of research is directed toward the discovery of methods of differentiating between degrees of intelligence, and, as we have seen, his attempts have ranged all the way from mental tests proper to a study of handwriting. The contributions to the *Année Psychologique* for 1905 are very largely the work of Binet alone, or of Binet and Simon in collaboration. First of all we have a short article on the question of mental fatigue; then an article criticizing Van Biervleit's proposal for the measurement of intelligence indirectly and simply by means of attention. There is also a practical article by Binet and Simon dealing with children who have been released from a school for the backward or feebleminded. They found in the institution investigated little knowledge of or interest in the children who had left. They urge the need of a psychological examination for all children in such institutions. This should be done yearly and incorporated in an annual report. They also deplore the lack of follow-up methods. This article is obviously an outgrowth of the work the two authors were doing in the feebleminded institutions.

There next appear in the volume three articles by Binet and Simon covering pages 163 to 336 inclusive. The first is theoretical and historical in tone. It brings out the unsatisfactory nature of our usual methods of diagnosing feeblemind-

edness, which are vague and subjective. It recounts all attempts up to that time at making the procedure more objective and the writers mention as praiseworthy the efforts of two French physicians, Blin and Damaye, in their attempt to draw up a list of questions with a method of scoring.

The First Scale.—The second article is entitled, "New Methods for the Diagnosis of the Intellectual Level of Abnormal Children." Here appears for the first time the idea of a scale of intelligence, *"une échelle métrique d'intelligence."* Here too we find the first specifications of intelligence tests, namely, they must be simple, must not consume a long time, they must be heterogeneous and not pedagogical. The thirty tests proposed are arranged in a series of increasing difficulty. They include many of the tests that we have noted in the previous work of Binet. The tests are not grouped according to age. They are merely to be scored with whole or half credit or no credit. So far the authors have not arrived at the idea of mental age. We have, however, in this article the original tests, the idea of a graded series, the concept of intelligence and a conception of the fundamental qualities of an intelligence test. This set of tests is sometimes called the 1905 scale.

In the third article, following immediately the one described above, we find the authors applying the tests they have just described to groups of normal and abnormal children. The article is very lengthy and minute details of application and the results obtained from children are given. It marks no further advance in test procedure. It merely gives us a picture of the authors at work in applying the tests to children. And the main thing that Binet and Simon seem to have gained from this experience is the necessity for laying more emphasis upon the methods of giving and scoring the tests. They are moving rapidly toward the concept of a standardization of procedure.

The 1908 Scale.—Little of importance for our topic was written by Binet during the next two years, although it is unquestionable that during this time he was trying out in actual practice the tests he had formulated in 1905. In 1908, there-

fore, appears his important article entitled "The Development of Intelligence in Children," in which he describes in detail the so-called 1908 Scale. This scale appears later in our book in Chapter VI, and a detailed description is, therefore, unnecessary here. The important point to note is that the tests are now grouped according to their appropriate ages. Furthermore the idea of a mental age is now introduced. This is one of Binet's important and valuable contributions to the problem of mental testing. The mental ability of any individual is expressed by the age that he reaches in the graded series of tests and this age is known as his mental age. To find out which tests were suitable to each age, Binet tested presumably normal children at each age and if from 60 to 90 per cent of such children passed the test, the test was considered standard for that age. Here, therefore, we see the beginnings of test standardization. The number of children tested and the method employed may seem unsatisfactory to us at the present time, but we must remember that this was pioneer work on the part of Binet and Simon.

Mental Age.—The use of mental age as a measure of intelligence is undoubtedly due to Binet, although the comparison of an individual in regard to knowledge or ability with a child of a specific age had often been made previous to the time of Binet. Woodrow (19) tells us that Esquirol in 1828, Duncan and Millard in 1866, and Down in 1887 had all made comparisons between feebleminded and normal children on the age basis without, however, employing any tests. Hall (48) in his description of the trial of one William Freeman, colored, charged with having killed four persons unknown to him and without provocation, reports that a Dr. Dimon, psychiatrist, said that the accused "in point of knowledge was equal to a child of two or three years." Undoubtedly many other references could be gathered of the use of mental age in this vague and haphazard manner.* It remained for Binet to take this

* See quotation by Gesell from Vol. 39, *Connecticut Reports*, Hartford, 1873. Gesell, A., *The Mental Growth of the Pre-School Child.* New York, 1925, 355-358.

relatively useless concept and crystallize it, to make it definite and concrete, and to raise it into one of the most practical and useful concepts for the psychologist and psychiatrist.

Miscellaneous Studies.—During this same year, 1908, Binet published several articles and they show his versatility and the wide range of his interests. One of these is a study of the results of a questionnaire among teachers of philosophy in French colleges and universities. He finds from this study that there is a growing dislike among students for formal logic and an increasing interest in science and sociology. Another article discusses the relation between language and thought. He compares an imbecile who has little speech with a real aphasic and questions Marie's dictum that every aphasic shows a diminution in his general intelligence. Language, he concludes, is not co-extensive with thought. He is inclined to hold to the theory of imageless thought. Still another article deals with palmistry in which the professional palmist is allowed to see only the hands of the subject and then to estimate the intelligence. Similar judgments are made from pictures of hands. The results are only a little better than pure chance. In a footnote to this article Binet excuses himself for dealing with palmistry by saying that he has investigated everything for a measure of intelligence, the head, the physiognomy, physical stigmata and now the hand. Thus we see Binet still casting around for measures of intelligence, while at the same time pushing forward with the scale of tests that was to achieve such signal success.

In 1909 we have a long and excellent article by Binet and Simon on the psychology of the feebleminded, covering a great many traits, such as attention, voluntary effort, movements in writing, the sense of pain, the number sense, and the like. There is much detail and several excellent descriptions and, as usual, the writers give the results of numerous tests to illustrate their points.

Insanity.—In the same year we see Binet applying his scale to particular types of insanity. He investigates the intelligence

of general paralytics and senile dements. The difference, he says, between the general paralytic and the imbecile is the difference between a disturbance in the *functioning* of the intelligence on the one hand and in the arrest of the *development* of the intelligence on the other. The scale is useful for it seems to show that previous knowledge existed in the general paralytics and the senile dements. Two other articles published this same year show again his diverse interests. One is an article on the psychology of painting and the other a psychological study of a particular painter's work.

In 1910 Binet summarizes all his attempts to diagnose intelligence from external physical signs in an article entitled, "The Physical Signs of Intelligence in Children." He recapitulates his work on the dimensions of the head, the so-called stigmata, the face and the hands, and sums up by saying that, although the averages of groups may in certain measurements show appreciable difference, we, nevertheless, cannot judge the *individual* child by these external signs. The same year we have, in collaboration with Simon, a long article attempting to define and describe the principal mental states of the various forms of insanity. This shows Binet's interest spreading over from feeblemindedness to insanity.

The 1911 Scale.—Binet's last important article on mental tests appeared in the *Année* for 1911, the year in which he died, and contained a further revision of his scale, the so-called 1911 Scale. This scale is given in Chapter VI of this book. It differs from the 1908 Scale in arrangement of tests and in the allotment of tests to each age. Some new tests are introduced and some of the old are dropped because they are too much like school work. Here we see the effect of a more exact differentiation between general intelligence and knowledge. He also raises the question of the relation of intelligence to school standing and replies to what seemed a criticism of the tests on the part of Decroly and Degand who had used his tests in Belgium. These workers had found that on the average the children tested were one and a half years advanced, and Binet

replies that it is perfectly feasible to suppose that such a difference really exists between the children of superior social standing tested by Decroly and Degand and the poor children of Paris on which his norms are based.

Tests for Soldiers.—There only remains to be mentioned a short article by Binet, written very probably shortly before his death, which is of interest in view of the unexpected development in mental testing in this country brought about by the World War. Binet summarizes in the *Année* an article written by himself and Simon in the *Annales Médico-psychologiques* for January 1910 on the need for a method of diagnosis to be applied to mentally defective soldiers. Binet and Simon seem to have taken up the matter with the military authorities and urged the adoption of psychological tests. They pointed out the desirability of eliminating the feeble-minded recruits. No progress was made because the medical officers thought the tests unsuitable. Seven years later the Binet-Simon tests, revised for American purposes, were used with splendid results along with other tests in helping this country build up a large and efficient army. We can well afford to give Binet credit for being the first to make such a suggestion, without detracting from the distinction merited by the American psychologists who began the work in the army and carried it to a succesful conclusion, and were not aware that this had been proposed by Binet himself.

Conclusion.—This brief survey of the writings of Binet shows extremely well how his attention and interest were from the start focussed upon the possibility of the measurement of intelligence. We have tried to show how he set this as his aim and how he worked incessantly at the problem, attacking it from every conceivable angle. The method finally adopted seems to us at the present time the obvious and natural method, but this was by no means the case at the time that Binet first started his work. His attempts to obtain an index of intelligence indirectly by means of head measurements or the two-point threshold on the skin reflects the dominant line of

thought in psychology at that time. It was natural that Binet should seize the tools that were ready to his hand and attempt to use them for his purposes. At the same time we notice that the intelligence test method is used by him from the start. It is crude at first but is gradually refined. And so we see him swinging over from the one method to the other, now trying the indirect approach through head measurements or aesthesiometry or graphology or palmistry, and now devising and developing a large number of intelligence tests as we know them to-day. Finally the intelligence test method wins and we see the Scale gradually emerging until it stands before us as a new method for the measurement of intelligence.

It is well to remember the gradual evolution of the Scale in the hands of Binet. The method is so simple that we are apt to forget that its growth was slow. The fact that the growth of the Scale was so slow and that every test was so well tried out, may be the reason why the Scale has fulfilled so well its initial promise. The original Binet-Simon Scale will probably soon disappear. It is in fact even now fast disappearing in the face of better revisions and new scales, but, nevertheless, all these are based upon the fundamental work accomplished by Binet. If, in the history of psychology, we call Wundt the father of experimental psychology, we must then call Binet the father of intelligence testing.

REFERENCES

Binet, A. See *L'année psychologique,* Vols. I to XVII, for most of the articles referred to in the text.

Binet, A. (98). "La mésure en psychologie individuelle." *Revue Philosophique,* Vol. XLVI, pp. 113-123.

Binet, A. (03). *L'étude experimentale de l'intelligence.* Schleicher, Paris.

Binet, A. and Simon, T. (16). "The Development of Intelligence in Children." Vineland Training School. Translation by E. S. Kite of several articles in the *Année.*

Binet, A. and Simon, T. (16). "The Intelligence of the Feebleminded." Vineland Training School. Translation by E. S. Kite of several articles in the *Année.*

Clarparède, E. (12). Alfred Binet. *Child Study,* Vol. V, pp. 18-24 and pp. 97-102. Translation by M. Louch.

Hall, B. F. (48). "The Trial of William Freeman." *Am. J. of Insanity,* Vol. V, pp. 34-60.

Martin, R. Alfred Binet. *Les Presses Universitaires de France.* Paris (no date).

Peterson, J. (25). *Early Conceptions and Tests of Intelligence.* Yonkers.

Woodrow, H. (19). *Brightness and Dullness in Children.* Philadelphia.

CHAPTER III

THE DEVELOPMENT OF INTELLIGENCE TESTING AFTER BINET

In this chapter we shall attempt to indicate the general development of intelligence testing following upon the work of Binet. We have noted in a previous chapter the many influences that were at work leading to the development of intelligence tests even before Binet, and these influences would naturally have borne fruit regardless of the great French psychologist. It was, however, Binet's work that gave a very definite turn to the course of this development, mainly in the direction of attempts to measure general all-round ability rather than specific abilities.

We shall make no attempt here to trace further the social aspect of our problem, as we did in the first chapter, where we attempted to give the general historical background for the rise of mental testing. We shall only mention such social institutions as were directly affected by the vigorous growth of intelligence testing in that branch of it which came to be known as clinical psychology. It is enough to remind the reader here that institutions for the feebleminded and defective delinquents, which we have previously noted as springing up here and there in the nineteenth century, have continued to increase in numbers and efficiency in the twentieth. Child-caring agencies of all kinds have greatly multiplied and special provision for backward or feebleminded children in the public schools is now taken as a matter of course in educational circles to-day. It would obviously take us too far afield to follow the development of these social agencies, all of which are making use of the intelligence tests devised by the psy-

chologist, and in turn presenting problems to him for solution, which have thereby enlarged the field of his research.

Without attempting to follow any strict chronological sequence, we may, nevertheless, distinguish certain definite phases in the recent history of intelligence testing. These phases overlap each other greatly in point of time, and it is useless to attempt any sharp line of demarcation. We may classify them as follows:—1. The Measurement Movement in Educational Subjects; 2. The Introduction of the Binet Scale into America; 3. The Controversy over the Validity of the Scale; 4. Clinical Psychology; 5. The Stanford Revision; 6. The Appearance of Other Scales; 7. The Group Test.

1. *The Measurement Movement in Educational Subjects.* —Simultaneous with the work of Binet in intelligence testing in France, there was arising in this country an interest in the objective measurement of school accomplishment. Rice's (97) report on spelling marked the beginning of the movement. But the real leader in the movement was Thorndike. Beginning in 1903 with his "Educational Psychology," we note the emphasis he lays upon measurement. In 1904 his "Introduction to the Theory of Mental and Social Measurements" appeared, a book which had a profound influence upon the measurement movement. Stone's "Arithmetic Tests," worked out under Thorndike's direction, were published in 1908, and in 1910 the Thorndike "Hand-writing Scale" appeared. From that time to the present, each year has seen the publication of one or more scales or tests for the measurement of educational subject-matter. Many of them have appeared under Thorndike's direction and all of them have been stimulated by his pioneer work. In addition to this Thorndike had done some work with mental tests, so that the introduction of the Binet Scale into America fell in line with the ideals and efforts of the measurement movement.

2. *The Introduction of the Binet Scale into America.*—Goddard was the first to introduce the Scale and to make use of it in this country. The psychological laboratory at the

Training School for Feebleminded Children at Vineland was founded in 1906, and Goddard was made Director of the laboratory. It is noteworthy as being one of the first laboratories in an institution for the feebleminded devoted primarily to psychological work. Goddard had seen something of Binet's work in France and was acquainted with his scale as it applied to feebleminded children. In 1908 he began to use the Scale at Vineland and was speedily convinced of its usefulness. He set to work adapting the Scale to American conditions, making as few changes as possible. Preliminary work with the Scale made him doubt whether all the tests were rightly placed for American children and he, therefore, took the next logical step in his standardization of the Scale on 2,000 American children. This standardization was published in 1910 and is based upon Binet's 1908 Scale. Goddard followed essentially Binet's method of standardization, although he made use of a much larger number of cases than Binet had used at any one time in the making of his scale. This revision of Goddard's remained for a long time the standard for American practice. The Training School for Feebleminded Children at Vineland, where Goddard was director of the psychological laboratory, became the Mecca of all those interested in the new clinical psychology, and the storm center around which much of the heated discussion as to the validity of the Binet-Simon Scale raged. The work at Vineland was ably supported by other men, notably by Kuhlmann and Huey, both psychologists attached to institutions for the feebleminded. Without going into detail we may say that the emphasis at first was laid upon the use of the Scale for the detection of feebleminded children. Vineland led the way in using the Scale as a means of classifying all the children in the institution, and very soon came to use it as a routine procedure for the admission of all new cases. From Vineland the use of the Scale spread rapidly to other institutions. The point to note, however, in the early use of the Scale is that for practical purposes its use was almost entirely restricted to the feebleminded.

Normal children were tested in the main solely for purposes of standardization. Some share of the prejudice against mental tests among the public at large must be attributed to this fact. To allow a child to be "submitted" to a mental test was, and still is, to some extent, equivalent to raising the question as to the integrity of his intelligence, because mental tests were from the beginning so closely associated with the feebleminded. This attitude towards mental tests is now rapidly disappearing, but it was a prejudice against which psychologists had to fight in their efforts to widen the scope of mental testing.

3. *The Controversy over the Validity of the Scale.*—From the very beginning the validity of the Binet Scale and the possibility of the measurement of general intelligence were challenged. Leaving aside criticism by non-psychologists, we find at the outset some psychologists who are opposed to the method as being unscientific and founded upon false premises. There were many who refused to recognize the work of the "mental testers" and regarded the whole matter as a fad that would soon fade into oblivion. As time went on, however, this attitude gradually disappeared, and no one at the present time would be bold enough to assert that mental measurement is a passing fad.

More serious and, sometimes, constructive criticism centered upon the actual tests themselves, the method of procedure and the problems of standardization. This criticism came from men who were more or less directly working in the field. It centered naturally around the Binet Scale. At the one extreme we find those who, while admitting the soundness of Binet's general conception, would, nevertheless, so revise and alter the tests and procedure as to leave practically nothing of the original Scale. At the other extreme are to be found those who treat the Scale as if it were something sacred and as if it had emerged perfect from the hands of the master. Between the two extremes are the workers who by actual trial and error modified, revised and enlarged the work that Binet began.

It might be amusing, but it could hardly be profitable to enter into a detailed account of this period of criticism. Much of it has no application at the present time, and it has already faded into the historic past so rapidly as to make one smile at the emotional outbursts and personal feelings that were aroused.

Much good resulted from these controversies nevertheless. The necessity for standardization of procedure, so that examiners could compare their results, was emphasized. A clearer line of demarcation was drawn between tests of intelligence and tests of knowledge. The need for other kinds of tests became obvious in cases where the Binet tests could not apply, especially in cases of language difficulty. The problems of standardization were discussed and the desirability of more careful placing of tests emphasized. All this discussion helped to a better understanding of just what "general intelligence" signified, and led us slowly to recognize certain conditions that a test ought to fulfill in order to be considered a good test of general ability. On the whole, therefore, the controversy that raged around the Binet Scale left us with clearer ideas as to the problems and methods of intelligence testing, even although much of it was mere opinion and useless.

4. *Clinical Psychology.*—Although clinical psychology proper dates back at least to the last decade of the nineteenth century, it is undoubtedly true that the Binet Scale was the one most potent factor in its development and expansion. Shortly after the first work with the Scale in the institutions for the feebleminded, we find psychological testing of all kinds spreading rapidly to juvenile courts, reformatories, prisons, children's homes and schools. The psychological clinic did not and does not depend upon the Binet Scale, but it is unquestionably true that the appearance of the Binet Scale acted as a tremendous stimulus to this type of work.

Notable among the juvenile court clinics was the one at Chicago under Healy and Bronner. Here no undue emphasis was given to the Binet Scale, but the value of the scale and

of other mental tests was well recognized in relation to the larger field of work covered by the clinic. Many other juvenile courts have found the services of clinical psychology of value and a few maintain special clinics, while others depend upon the help of outside psychologists.

Among clinics established in reformatories and other penal institutions, one of the most noteworthy was the psychological clinic at the New York State Reformatory for Women at Bedford Hills. Real constructive work in the way of enlarging the field of intelligence testing has been done here under the direction of psychologists. Many other reformatories have had or still have clinics, where the scope of the work ranges from a very exhaustive examination down to the giving of a short set of tests. Many reformatories, prisons and penitentiaries have been surveyed by psychologists and their results have been of value in increasing our knowledge of the mentality of the criminal.

And, lastly, many school systems have recognized the importance of psychological work. School clinics have been established in connection with the education of backward and feebleminded children. As we have seen in Chapter I, especial attention to the backward child began in the nineteenth century. At first the interest was entirely pedagogical, but it was not long before the help that psychology could render was appreciated and we see the coming of the school psychologist, whose duty it is to select the backward children for the special class. The number of these clinics, or departments of child study, is now great, and the work they are doing varies from minute and thorough psychological testing down to the mere giving of Binet tests to aid in the segregation of the mentally retarded. There is a place in every school system for psychological tests and unquestionably the future will see a great expansion of their use in our schools. At the present time the center of interest in psychological work in the schools has shifted from the mere segregation of the mentally retarded to a wider use of intelligence tests in the classification of

pupils in general and much attention is being paid to the child of superior intelligence. We are beginning to realize that the bright child as well as the dull needs special educational guidance.

The rapid expansion of the use of intelligence tests in our schools has assisted and been assisted by the development of educational tests and measurements. The rise of such measurements forms an interesting chapter in school history but it lies outside the scope of this work. Undoubtedly the two kinds of measurement, the one probing the native ability, the other the knowledge of the child, are complementary, and the future will see a more intimate use of the two combined.

5. *The Stanford Revision.*—The publication by Terman of the Stanford Revision of the Binet-Simon Scale in 1916 marks a distinct advance in intelligence measurement. Even before the publication of the Binet-Simon tests, Terman had been interested in the problem of individual differences in intelligence among school children, and shortly after the first publication of the Binet tests by Goddard in 1908, he seems to have become interested in Binet's method. During 1910-11 Terman and Childs constructed a tentative revision and extension of the Binet 1908 Scale, which they published in 1912. Terman considered this merely a tentative revision, because his experience with the Scale so far had shown him the great possibilities in the way of further extension and more complete standardization. This revision, extension, and standardization occupied Terman and his co-workers for the next five years and the final result certainly justified the labor and time involved.

The Stanford Revision adds nothing essentially new to the ideas of Binet. What Terman did, however, was to work out more thoroughly and more accurately the method suggested by Binet. The scope of the standardization was broadened and the scale was so adjusted as to fit accurately at each age. Much of what is only implicit in Binet is made explicit by Terman. One other important contribution by Terman is

worth noting, namely, the adoption of the Intelligence Quotient as suggested by Stern.* The use of the I.Q. by Terman in connection with his Scale brought this measure of intelligence into common use.

6. *The Appearance of Other Scales.*—It was natural that the stimulus given to mental testing by the use of the Binet-Simon Scale should result in the construction of other scales for measuring intelligence. If there are few, it is not because of a scarcity of ideas as to different tests, but rather because of the labor involved in standardization. We shall mention here the more or less well standardized scales, which are not obvious revisions of the Binet, but which have been stimulated by it. The Point Scale by Yerkes, Bridges and Hardwick (15) used the original Binet Tests but substituted a scoring method for each test and a total score instead of the "all or none" method and the age grouping of tests in the Binet.

The Pintner-Paterson Performance Scale (17) was the first scale to use tests entirely different from those of Binet. None of the tests in the scale require a knowledge of language by the subject. In this respect also it is very different from the Binet, where directions are given orally and most of the responses are oral. The scale has been found useful in the testing of foreign children, deaf children and also as a supplement to the Binet Scale. The Army Performance Scale (Memoirs 21) was constructed to supply the army examiners with a means for measuring the mentality of the foreigners in the army when an individual examination, in addition to a group test, was required. Several of the tests used in the Pintner-Paterson Performance Scale are included in this scale, in addition to many others. The scale has not been standardized on children, but the scores can be turned into an approximate mental age based upon a comparison of Performance Scale and Stanford and Point Scale records gathered in the army.

* Stern, W. "Der intelligenz quotient als mass der kindlichen intelligenz." *Zsch. f. angew. Psychol.*, 1916, 11, 1-17. In this article he refers to his proposal in 1912 at the Fifth Congress of Experimental Psychology with reference to the use of the intelligence quotient.

The Herring Scale (22) follows the Binet in type of test, but all the tests used are different. It also differs in method of scoring and of computing mental age. It is intended to cover the same ground as the Stanford Revision and it correlates highly with the latter.

7. *The Group Test.*—The most recent development in mental testing has been the rise of the group test. Here the subjects are examined in groups in contradistinction to the tests we have just described, in which the examiner is engaged with only one individual at a time. There is, of course, nothing intrinsically new in the group method, because it has been employed for a long time in psychological experiments. For the most part, however, these experiments were not directly concerned with getting an intelligence rating of the individual, and the new turn given to group tests at the present time is precisely the attempt to obtain a measure of the individual's mental ability.

For a long time there existed a distinct prejudice on the part of psychologists against the group test. Although the method had never been adequately tried out all sorts of *a priori* reasons as to why it could never be feasible were advanced, and most of these reasons have disappeared into thin air now that the method has been tried and developed. The group method does not give us all the information gleaned by the expert clinician in his individual examination, but it can give a very reliable intelligence rating.

We see the beginnings of the group method in the work of Thorndike and others with the early mental tests. After the introduction of the Binet tests, it was natural that the quicker group method should be tried out in order to arrive at a mental rating of more than one child at a time. Thus Pyle (13) gave several tests to groups of children, but did not combine the results of all the tests into one index to denote a child's intelligence. This was done by Pintner. Influenced by the work of Pyle and confronted with the practical problem of having to weed out the feebleminded from among a large group

of children, Pintner (17) used a set of six tests as group tests and took the median of the six percentiles as a measure of the child's mentality.

At the same time group tests were being used by Scott in his work of measuring the ability of business executives and salesmen. From these and other beginnings it is obvious that group testing was in the air and that it would have developed slowly toward the goal that it has now attained. The entry of this country into the World War and the consequent utilization of intelligence tests in the army accelerated enormously the development of the group test. The different tests at present available will be discussed fully in Chapter VII.

Summary. Considering the broad outlines of the development of intelligence testing since the work of Binet, we may discern three phases. The first is the introduction of the Binet Scale into this country, leading to minute work and much discussion of the principles involved, stimulating the development of clinical psychology and of intelligence testing in schools and institutions of all kinds and culminating with the Stanford and Herring Revisions, as representing the latest and most modern development of Binet's original idea.

Secondly, we see the construction of other kinds of Scales, diverging more or less from the principles of the Binet Scale, broadening our conception of the intelligence scale and extending the field of intelligence testing to different types of subjects.

And, thirdly, we note the rapid rise of the group test, introducing the concept of intelligence surveys of large numbers of individuals, breaking away from the method of the Binet Scale and opening up unlimited possibilities for the psychologist in the future.

REFERENCES

Goddard, H. H. (10). "The Binet-Simon Measuring Scale for Intelligence." *The Training School,* Jan.

Herring, J. P. (22). *Herring Revision of the Binet-Simon Tests.* Yonkers.

Memoirs, (21). "The National Academy of Sciences," Vol. XV, Part I, Chap. 5.

Pintner, R. (17). "The Mentality of the Dependent Child, together with a Plan for a Mental Survey of an Institution." *J. of Ed. Psych.,* Vol. VIII, April, pp. 76-88.

Pintner, R. and Paterson, D. G. (17). *A Scale of Performance Tests.* New York.

Pyle, W. H. (13). *The Examination of School Children.* New York.

Rice, J. M. (97). "The Futility of the Spelling Grind." *Forum,* Vol. XXIII, pp. 163-172; 409-419.

Stone, C. W. (08). *Arithmetical Abilities and Some Factors Determining Them.* T. C. Contribs. to Ed. No. 10.

Terman, L. M. (16). *The Measurement of Intelligence.* Boston.

Terman, L. M. and Childs, A. G. (12). "A Tentative Revision and Extension of the Binet-Simon Measuring Scale of Intelligence." *J. of Ed. Psych.,* Vol. III, pp. 61-74; 133-143; 198-208; 277-289.

Thorndike, E. L. (03). *Educational Psychology.* First Edition.

Thorndike, E. L. (04). *An Introduction to the Theory of Mental and Social Measurements.* T. C. Columbia University, New York.

Thorndike, E. L. (10). "Handwriting." *T. C. Record,* Vol. XI, No. 2.

Yerkes, R. M., Bridges, J. W. and Hardwick, R. S. (15). *A Point Scale for Measuring Mental Ability.* Baltimore.

CHAPTER IV

THE CONCEPT OF GENERAL INTELLIGENCE

Gradual Growth of the Concept.—In the preceding chapters we have been speaking of the measurement of general intelligence, without having attempted in any way to define the meaning of this term, and it is not easy even at present to do so. As a matter of fact general intelligence has rightly been assumed to exist and psychologists have gone about the measurement of an individual's general ability without waiting for an adequate psychological definition. This must not be taken to imply, however, that the psychologist was blindly measuring something, without having any conception of what he was trying to measure. In every case the psychologist has some rough working hypothesis upon the basis of which he proceeds, and it is these hypotheses and the logical conclusions that follow from them that we shall now briefly survey.

As we have noted in our historical account, the earlier attempts at mental measurement were concerned with the measurement of separate faculties, processes or abilities. Tests were devised to test different kinds of memory, attention, imagination and the like. Binet's efforts were at first largely devoted to the measurement of the higher and more complex processes such as reasoning, imagination and so on. Again Binet very frequently compared intelligent and unintelligent children in their methods of reacting to certain tests. Note carefully that his intelligent and unintelligent children were not so called on the basis of psychological tests, but simply on the basis of the ordinary subjective judgment of the teacher or parent, i.e., the ordinary common-sense judgment of the world. It is easy to imagine, therefore, that Binet became almost unconsciously

familiar with intelligence as a complex trait of an individual's make-up and very probably began to use unconsciously the type of reaction he had found common in his subjects as a criterion by which he might judge intelligence or lack of intelligence in unknown subjects. Be that as it may, the transition from his early work on groups of intelligent and unintelligent subjects, as diagnosed by the world, soon led him consciously to reverse the process and to diagnose the intelligence of his subjects by means of tests.

It is well to insist upon the fact that the meaning of general intelligence has had a gradual growth and that we did not start with a clear definition of general intelligence set up by some psychologists and forthwith proceed to measure it by means of tests. We may say rather that the psychologist borrowed from every-day life a vague term implying all-round ability and knowledge, and in the process of trying to measure this trait he has been and still is attempting to define it more sharply and endow it with a stricter scientific connotation.

As we go through the writings of Binet we do not find any well-formulated and standard definition of general intelligence, but rather numerous descriptions of it, now emphasizing one of its aspects and now another. For a long time, indeed, he did not seem to differentiate it clearly from knowledge, just as it is not at present clearly differentiated from knowledge in popular thought. Later on, this differentiation from knowledge creeps in, but only after he has had considerable experience with measuring intelligence by means of his scale of tests. Practical work with the scale made this differentiation necessary. This is a good example of the manner in which the meaning of the term "general intelligence" is acquiring a definite psychological connotation.

Binet considered attention and adaptation as the two most important factors in general intelligence, although we must not forget that he employs the term attention in a much wider sense than is ordinarily the case in psychology. One of his longest articles is devoted to a study of attention and adapta-

tion in order to see how intelligent and unintelligent children differ in these two respects, and many of the tests he devised to bring out these differences were later found useful in his scale. Judgment, again, is a term Binet (16) is fond of using and emphasizing, in his concept of intelligence. Intelligence, he says, "is judgment, or common sense, initiative, the ability to adapt oneself." And again, "To judge well, understand well, reason well—these are the essentials of intelligence." "Binet's definition emphasized three phases of behavior: (1) the ability to take and maintain a given mental set; (2) the capacity to make adaptations for the purpose of attaining a desired end; and (3) the power of auto-criticism" (Hollingworth, 20). Descriptions of this kind are Binet's contribution to a definition of general intelligence.

Definitions.—Many definitions of intelligence have been proposed by psychologists, and these vary greatly according to the standpoint from which the psychologist views this problem. Although they differ greatly, they are not therefore contradictory. We may group the numerous definitions proposed into four groups which we shall label: (1) biological; (2) educational; (3) a faculty; (4) empirical.

(1) *Biological.* Here the emphasis of the definition is upon the adjustment or adaptation of the organism to its environment or to certain aspects of its environment. We may conceive of intelligence as being the determining factor in the evolutionary process. The more intelligent organisms can adjust to a greater number of environmental changes. The "higher" the organism, the greater is its adaptability and the greater is its intelligence.

Stern's (14) definition of intelligence belongs here. He says, "Intelligence is a general capacity of an individual consciously to adjust his thinking to new requirements." And again, and perhaps better, "It is general mental adaptability to new problems and conditions of life." By this definition Stern differentiates general intelligence from talent which is the development of a specific ability, and again from knowledge or mere infor-

mation, and lastly from memory because this deals with the old, whereas general intelligence is directed toward the new. Further he says, "Any sort of perceptive, memorial or attentive activity is at the same time an intelligent activity just in so far as it includes a new adjustment to new demands." These statements of Stern are often further contracted as follows: "General intelligence is the ability of the organism to adjust itself adequately to new situations."

Similar to Stern's definition is the statement of Wells (17): "Intelligence means precisely the property of so recombining our behavior-patterns as to act better in novel situations." And Peterson (Symposium, 21) says, "Intelligence seems to be a biological mechanism by which the effects of a complexity of stimuli are brought together and given a somewhat unified effect in behavior." Woodworth (21) describes what the subject in a test must do:—"He has to see the point of the problem now set him, and to adapt what he has learned to this novel situation." Edwards (28) defines intelligence as "capacity for variability or versatility of response."

All of these definitions conceive of general intelligence as including behavior that leads to better and better adaptation not only in man, but in the whole animal kingdom. It includes the capacity for getting along well in all sorts of situations. This implies ease and rapidity in making adjustments and, hence, ease in breaking old habits and in forming new ones. Fundamentally, this leads us back to the general modifiability of the nervous system. An organism whose nervous system is very modifiable, can adjust itself to new types of situations and react adequately to a great number of different situations. The intelligent organism has a multiplicity of responses; the unintelligent few. The intelligent organism responds to a great number of situations; the unintelligent to few. Intelligent behavior leads one from one thing to another in ever-widening circles; unintelligent behavior is narrow and restricted, and leads to repetition or cessation.

(2) *Educational.* This kind of definition places the emphasis upon learning ability. As Buckingham (Symposium, 21) says, "Intelligence is the ability to learn." In a similar way Colvin (Symposium, 21) writes, "An individual possesses intelligence in so far as he has learned, or can learn, to adjust himself to his environment." Perhaps we might also include here Henmon's (Symposium, 21) statement that intelligence is the capacity for knowledge and the knowledge possessed. The knowledge possessed is not ordinarily included in intelligence by most writers.

This educational point of view with regard to intelligence stresses the educability of the individual, it emphasizes his learning power. That individual is intelligent who learns readily and easily. The individual who has difficulty in learning, who finds learning hard, is lacking in intelligence. There is no contradiction here with the biological point-of-view, because all learning may be regarded as adjustment or adaptation to various situations.

(3) *A Faculty.* Here intelligence is described in terms of a faculty or a capacity. The definition is not so much in terms of what intelligence does, but rather in terms of what intelligence is. The attempt is generally to delimit or restrict intelligence and set it off from other powers or faculties of the mind. Binet's various definitions belong mainly here, as when he says intelligence is judgment or common sense. Terman (Symposium, 21) makes intelligence synonymous with abstract thinking in his statement, "An individual is intelligent in proportion as he is able to carry on abstract thinking." Woodrow (Symposium, 21) calls intelligence an acquiring capacity. It is a capacity to acquire capacity. Again Haggerty (Symposium, 21) says, "It is a practical concept connoting a group of complex mental processes traditionally defined in systematic psychologies as sensation, perception, association, memory, imagination, discrimination, judgment and reasoning. . . . For the most part I would exclude from

the concept, emotion, instincts, will-activities and so-called character traits."

(4) *Empirical.* Here belong a certain number of definitions which emphasize the practical results of intelligence. They might also be called behavioristic, in the sense that they call attention to the behavior value of intelligent responses. Thus Thorndike (Symposium, 21) says, "We may then define intellect in general as the power of good responses from the point of view of truth or fact." And Ballard (22) would define intelligence as "the relative general efficiency of minds measured under similar conditions of knowledge, interest, and habituation." Pintner (26) writes, "We must get away from the idea that certain stimuli constitute in and of themselves adequate intelligence tests on all occasions. We must free ourselves from the idea that there is a specific faculty of intelligence. We must remember that intelligence is merely an evaluation of the efficiency of a reaction or group of reactions under specific circumstances." And Piéron (26) in much the same way remarks that, "intelligence does not exist in the mental mechanism; it is only an effect, a functional resultant under certain defined conditions, a behavior value."

There are other definitions of intelligence which can hardly be classified under any of the above four headings. Thus Thurstone (24) says that, "Intelligence is the capacity to live a trial-and-error existence with alternatives that are as yet only incomplete conduct." Unintelligent action is impulsive. The more we inhibit, the further back we push inhibition, the more chance is there for intelligent action. "Intelligence, considered as a mental trait, is the capacity to make impulses focal at their early, unfinished stage of formation." And Freeman (26) writes: "Psychologically, degrees of intelligence seem to depend on the facility with which the subject-matter of experience can be organized into new patterns. This rearrangement of thought material is what characterizes particularly the higher mental processes."

Such diversity of opinion makes it impossible to give the

reader any one definition of intelligence and say that this is what psychologists think intelligence is. But as a matter of fact the differences between psychologists are more apparent than real. They spring from different points of view. There is inherently no difference between the biological and educational point of view. They both depend upon the modifiability of the organism. The faculty point of view tries to specify what particular processes or powers are used in making the adaptations and adjustments and learnings, whereas the empirical point of view calls our attention to the result of the responses of the organism. According as this result is good or poor, we infer intelligence or the lack of it in the responding organism.

Description.—Let us now go on from these brief definitions to a more elaborate description, and by describing fill in our concept of intelligence. Intelligence has certain properties or characteristics, and these we shall call attributes. Again, intelligence may be manifested in different situations and so we shall divide intelligence into different kinds.

(1) *Attributes.* Thorndike (27) distinguishes four attributes or aspects of intelligence, namely, (a) level; (b) range; (c) area; (d) speed.

(a) Level. This refers to the degree of difficulty of a task that can be solved. If we think of all tasks or test items or situations as varying in difficulty from very easy to very hard, and if they are arranged on a ladder or scale of difficulty, then the height that we can attain on this ladder determines our level or altitude of intelligence. How difficult a problem can we solve; how hard a task can we master? This will determine our level of intelligence. It is more difficult to interpret a picture than to describe one, and more difficult to describe than to enumerate the things in a picture, and hence enumeration, description and interpretation are found at three different levels on the Binet Scale.

Level or altitude is the most important aspect of intellect, but we can never measure it alone. The other necessary

aspects or attributes of intellect are always present. We cannot measure height without some kind of width along which to measure.

(b) Range. Range or width refers to the number of tasks at any given degree of difficulty that we can solve. If all the tasks or situations in the world could be accurately measured with respect to the amount of intelligence required to solve or meet them, then we could group these tasks at each level of intelligence, and the number of such tasks which any individual could solve at any given level would represent the range of intelligence for that individual at that level. Theoretically an individual possessing a given level of intelligence should be able to solve the whole range of tasks at that level. Practically he is not able to do so because of lack of opportunity or lack of knowledge of other facts at lower levels which are prerequisites to the solution of tasks at higher levels. Range then is determined not only by level but also by breadth of experience, by opportunity to learn. In our intelligence tests range is represented by items of equal difficulty. Range is positively correlated with level. Thorndike (27) finds actual correlations ranging from .96 to .99. Given the same opportunity to learn individuals of higher levels of altitude possess greater ranges at all levels than those of lower levels of altitude. This seems to be due to a better all round memory and a greater speed of learning in the individuals of higher levels. An imbecile can master all the items at a given low level of intelligence, but he is slow in learning them and by the time he has learned to respond to the thousandth task, he has forgotten the responses to the first ten. The intellect of higher level learns all the thousand, learns them quickly and does not forget any. And so it would seem to be with intellects of all levels and hence the positive, but not perfect, correlation between altitude and range.

When we talk about intelligence as being inborn and inherited, we are really referring to altitude of intellect. We are discussing an abstraction that does not exist, because

in actual life we can only estimate the altitude by means of test items and the test items always indicate a range. Our tests must always be made up of so many items at different levels. We cannot measure altitude without range or width. In comparing individuals with reference to their intelligence we must be sure that the items on which we make such comparisons fall within the range suitable to the individuals compared. Children who attend school cannot be compared with children who do not attend school by means of tests which presuppose the ordinary school environment, which so many tests do.

(c) Area. Area may be thought of as the total number of situations at each level to which the individual is able to respond. Area is a summation of all the ranges at each level of intelligence possessed by an individual. It is in general very highly correlated with altitude. Area is not of great practical importance because we cannot hope to measure by means of tests the total area of any one's intelligence. Theoretically this attribute of area will, however, help us in our concept of intelligence. We can readily conceive of individuals of equal altitude of intelligence differing in area. Let Figure 1 represent the area of two individuals of equal altitude. The height of the figure represents the altitude. Both R and S reach altitude F. They are both able to do tasks of F difficulty. Let the horizontal lines at A, B, C, D, E, and F represent the range or width at each level. Then it is seen that at all levels R has a wider range than S, and, therefore, the total area of intellect of R is greater than that of S. This difference in area may be due to differences in opportunity, in interest, in zeal or temperament.

The altitude of intellect of a rural child in a narrow environment may be the same as that of a New York youngster, but the area will probably be smaller. The altitude of an imbecile in a state institution may be the same as that of an imbecile in a wealthy family surrounded by tutors and governesses, but the area of the former, if we could adequately measure it,

would undoubtedly be smaller than that of the latter. Many a learned man of vast scholarship has no higher an altitude of intellect than a narrow and limited practical man.

(d) Speed. This is the rapidity with which we can respond to situations. Speed shows a positive correlation with altitude. Speed of correct performance of easy tasks correlates about .50 with altitude (Hunsicker, 25). Speed is much less closely bound up with altitude than are the other attributes.

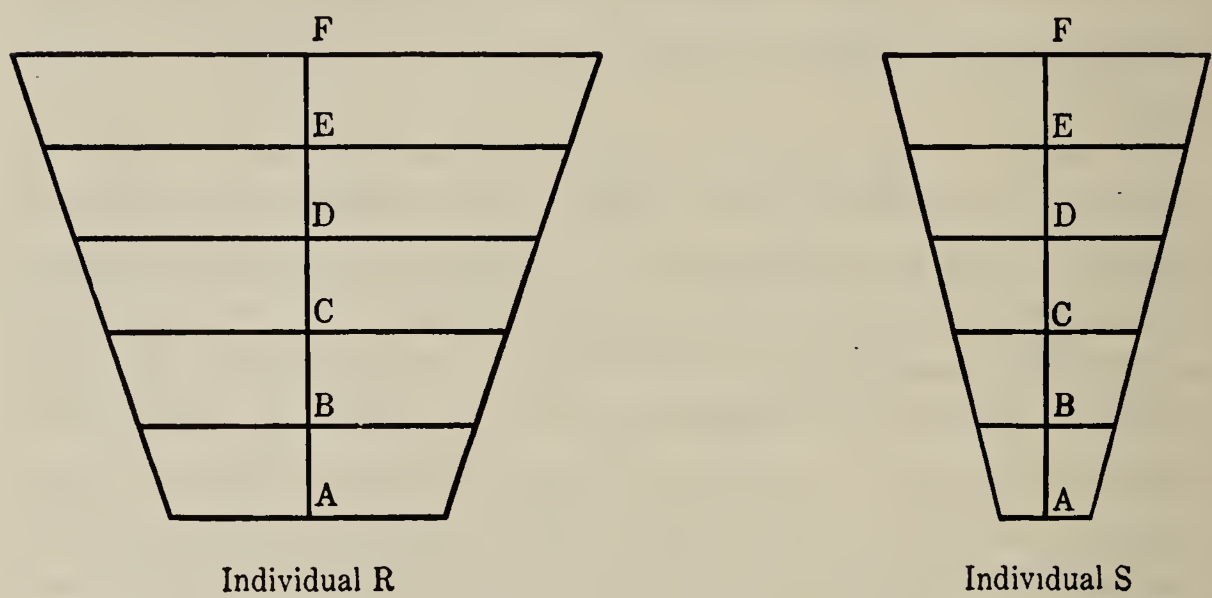


Fig. 1.—The Attributes of Intelligence.

We should not, therefore, emphasize speed too much in our intelligence tests.

The usual intelligence rating based upon a Binet Test or upon one of the common group intelligence tests is necessarily a mixture of all these attributes. Necessarily, because they never can be separated in practice. When we test a person we must give him a certain number of tasks (area), and these tasks vary in difficulty (altitude), and there are a certain number at each level of difficulty (range), and they are responded to in a given time (speed). But some tests emphasize one aspect rather than another. Thus the Binet emphasizes level and range and minimizes speed. The Army Alpha lays more emphasis upon speed. The Thorndike CAVD emphasizes altitude and pays practically no attention to speed. If we

combine the results of several intelligence tests of the same individual, we increase the area of his measured intelligence.

(2) *Kinds.* In addition to these attributes or characteristics of intelligence we may also think of different kinds of intelligence. Thorndike has suggested a three-fold division into (a) abstract; (b) concrete; and (c) social.

(a) Abstract. Here we test the ability to respond to symbols of various sorts, such as words, numbers, letters and the like. In this sense all of our verbal tests are tests of abstract intelligence. Abstract intelligence is required for the ordinary academic subjects in school, such as reading, arithmetic, history, geography and so on. At its highest levels abstract intelligence is seen in the reactions of the student and philosopher dealing with the relations of things symbolized in words or numbers or mathematical formulæ.

(b) Concrete. Here we test the ability to respond to things themselves rather than to the verbal symbols of them. Concrete intelligence is not synonymous with manual dexterity or mechanical ability, but it is the ability to comprehend actual concrete situations and react adequately to them. In everyday life such ability is shown in our reactions to locks and keys, cupboards, electric lights and switches, automobiles, trains, elevators, lawn-mowers and the innumerable utensils, machines and instruments that people in different walks of life handle every day. In its higher levels, it is typified by the inventor, the surgeon, the musician and others, in so far as they are reacting to concrete things. This kind of intelligence is best measured by means of performance tests. Indirectly it may also be measured by picture tests.

(c) Social. Here we are thinking of the ability to react to persons as distinct from concrete things and abstract symbols. Social intelligence would not include the feelings or emotions aroused in us by other people, but merely our ability to understand others and to react in such a way towards them that the ends desired should be attained. High social intelligence is presumably possessed by those who are able to

handle people well. Good administrators, teachers, politicians, preachers, generals and executives of all kinds should possess high social intelligence. As yet we have no adequate tests of social intelligence.

This three-fold division of intelligence is merely a convenient scheme for dividing the numerous reactions which may be indicative of intelligence. There is no reason why other divisions should not be made, and, indeed, it is quite common to speak of verbal as opposed to non-verbal intelligence, and when we do this we are cutting across the previous scheme of classification and considering all reactions as divisible into verbal and non-verbal. Or, again, we may take abstract intelligence and divide it into verbal and numerical intelligence. Or, we may with Thorndike divide abstract intelligence or all intelligence into as many different kinds as there are different test situations, and so we may speak of C intelligence, meaning thereby intelligence as shown in verbal completion, and of A intelligence, meaning thereby intelligence as shown in dealing with arithmetical situations, and of V intelligence which refers to vocabulary, and of D intelligence which means intelligence in following verbal directions. And then we may talk of CAVD intelligence which is, of course, a combination of the four particular kinds of intelligence just described.

If we go to the actual results of testing and ask whether we find there any warrant for our three-fold classification of kinds of intelligence we find such results as are indicated in Table I.

This table shows samples of correlations between various kinds of tests. Section I shows the results for what we have called abstract intelligence tests correlated with other abstract tests. Under abstract have been included the Binet tests and the usual verbal group tests. The Binet tests are not entirely abstract but they are very largely so. Under Section II we have samples of correlation between abstract and concrete tests. Under concrete there have been included performance tests, non-language pencil and paper tests, mechanical tests

TABLE I

Correlations Between Different Kinds of Intelligence

I. Abstract with Abstract

Author and Date	*Tests Correlated*	*r*	*n*	*Subjects*
Root (22)	Terman Group with other verbal group tests. Av. of 3 correlations	.73	21	Grade IX
Avery (23)	Binet with Terman Group	.80	155	Age 11
Haggerty (23)	Haggerty Delta 2 with N.I.T.	.81	500	Grades IV-VIII
Jordan (23)	Stanford Binet with 4 Group Tests separately. Av. of 4 correlations	.64	64	High School pupils
Herring (24)	Stanford Revision with Herring Revision	.98	270	Ages 4-18
Miller (24)	Miller Test with 9 verbal group tests combined	.89	57	1st year high school
Flemming (25)	Terman Group with Miller Group	.78	60	High School seniors
Garrison and Robinson (25)	Stanford Binet with N.I.T.	.72	131	Grades III-VIII
Tilton (26)	Six verbal tests. Av. of 9 correlations	.53	250	Ages 11-17

Author and Date	*Tests Correlated*	*r*	*n*	*Subjects*
Wentworth (26)	Stanford Binet with Dearborn A. Av. of 3 correlations	.71	575	Grade I
Pintner (27)	Terman Group with Pintner Rapid Survey	.82	138	Age 11
Walters (27)	Stanford Binet with N.I.T. and Otis. Av. of 3 correlations	.63	165	Grades VI-VII
Dougherty (28)	Binet with Pintner-Cunningham	.80	50	Kindergarten
Sangren (29)	Pintner-Cunningham with Stanford Binet	.68	100	Grade I
	Pintner-Cunningham with 7 group tests	.75	100	Grade I

II. Abstract with Concrete

Author and Date	*Tests Correlated*	*r*	*n*	*Subjects*
Yerkes (21)	Binet M.A. with Cube Construction	.63	260	Adults
Franzen (22)	Thirteen verbal group tests with Myers Mental Measure. Av. correlation	.46	57	First year high school

Morgenthau (22)	Binet with Pintner Non-Language	.44	97	Ages 7-16
Kohs (23)	Binet M.A. with Block Design M.A.	.82	366	Ages 5-20
Stenquist (23)	Verbal Group Tests with Stenquist Mechanical	.40	100	Grades VII-VIII
Pintner (24)	Various Verbal Group Tests with Pintner Non-Language	.25 to .72	22 to 620	Various Ages
Gaw (25)	Binet I.Q. with Performance Scale I.Q.	.41 to .49	100	Age 13
Johnson (25)	Binet M.A. with Pintner-Paterson Performance Scale	.83	488	Ages 3-13
Brown (26)	Haggerty Delta 2 with Pintner Non-Language	.43 to .59	456	Ages 10-13
Goodenough (26)	Binet I.Q. with Goodenough Drawing I.Q.	.74	334	Ages 4-10
Worthington (26)	Binet with Single Performance Tests	.41 to .79	Over 100	Clinic children, all ages

Author and Date	*Tests Correlated*	*r*	*n*	*Subjects*
Bronner (27)	Binet with Healy Pictorial Completion II	.44	282	Age 14
Brooks (27)	Binet plus 9 Group Tests with Pintner Non-Language	.46	108	Grade VII
Jones (27)	Multimental with Myers Mental Measure	.43	278	Ages 11-12
Walters (27)	Binet M.A. with Pintner Non-Language	.13	158	Grades VI-VII
Bell (28)	N.I.T. with Stenquist	.25	72	Grade V

III. Abstract with Social

Author and Date	*Tests Correlated*	*r*	*n*	*Subjects*
Broom (28)	Thorndike Intelligence with Moss Social Intelligence Test	.56	258	Students
Hunt (28)	Group Verbal with Moss Social Intelligence. Av. of 5 correlations	.47	689	Students
Pintner and Upshall (28)	Group Verbal with Moss Social Intelligence	.68	33	Students
Garrett and Kellogg (28)	Thorndike Intelligence with Moss Social Intelligence	.42	118	Students

Strang (30)	Group Verbal with Moss Social Intelligence	.44	311	Students

IV. Concrete with Social

Author and Date	*Tests Correlated*	*r*	*n*	*Subjects*
Hunt (28)	Two mechanical aptitude tests with Moss Social. Av.	.17	256	High School and College

and a drawing test. Under social we have, up to the present time, only one test which claims to measure this kind of intelligence and the results are shown under Sections III and IV of the table.

The average of the thirteen correlations of abstract with abstract is about .75. The average of the correlations between abstract and concrete is about .50. It would seem as if these tests of concrete intelligence were measuring something slightly different from the abstract tests. If they are valid tests of intelligence, they are measuring a different kind of intelligence. With reference to social intelligence the results of actual testing are at present too few to show anything of value.

Remembering, then, that any classification into kinds of intelligence is merely for our own convenience, we may, nevertheless, hold to our three-fold classification as useful. At the present time the great majority of tests may be thought of as measuring abstract intelligence. This is particularly true of our group tests. The Binet-Simon Tests are also largely abstract in type. Obviously abstract intelligence is very important and is closely related to school work. It is more easily measured. Concrete intelligence is measured by our performance scales and to some extent by tests involving pictures rather than words. Social intelligence is not yet adequately measured.

Theories of General Intelligence.—So far we have been dealing with definitions or descriptions of general intelligence, which make no attempt to explain in a more thorough-going manner the nature of general intelligence or its relation to the different abilities or traits of the individual. There are, however, one or two theories of intelligence that attempt to go further than the mere descriptions.

1. *Spearman's Theory.* Spearman's (27) theory of the nature of general intelligence is the one that has been formulated in greatest detail and it merits particular attention. Spear-

man * found on examining the correlations between various tests that these correlations "tended towards a peculiar arrangement which could be expressed in a definite mathematical formula." This formula

$$r_{ap} \times r_{bq} - r_{aq} \times r_{bp} = 0$$

is called the tetrad equation and the value constituting the left side of it is the tetrad difference.

We quote from Spearman (27): "An illustration may be afforded by the following imaginary correlations between mental tests (actually observed correlations will be given in abundance later on):

		Opposites	*Completion*	*Memory*	*Discrimination*	*Cancellation*
Opposites	1	..	.80	.60	.30	.30
Completion	2	.80	..	.48	.24	.24
Memory	3	.60	.48	..	.18	.18
Discrimination ...	4	.30	.24	.18	..	.09
Cancellation	5	.30	.24	.18	.09	..

For instance, let us try the effect of making

a denote Opposites.
b denote Discrimination.
p denote Completion.
q denote Cancellation.

From the table of correlations above, we see that r_{ap} will mean the correlation between Opposites and Completion, which is .80. Obtaining in a similar fashion the other three correlations needed, the whole tetrad equation becomes—

$$.80 \times .09 - .30 \times .24 = 0$$

which is obviously correct. And so will be found any other application whatever of the tetrad equation to this table.

* Quotations from Spearman, C. "The Abilities of Man." By permission of the Macmillan Company, publishers.

"The Two Factors. So far, the business is confined to matters of observation; we simply try out the tetrad equation on any table of actually observed correlations and examine whether it fits. The next step, however, is not observational, but purely mathematical; we have to ask how, if at all, this equation between the correlations bears upon the individual measurements of the correlated abilities. The answer is that there has been shown to exist a very remarkable bearing indeed. It is to the effect, that, whenever the tetrad equation holds throughout any table of correlations, and only when it does so, then every individual measurement of every ability (or of any other variable that enters into the table) can be divided into two independent parts which possess the following momentous properties. The one part has been called the 'general factor' and denoted by the letter *g;* it is so named because, although varying freely from individual to individual, it remains the same for any one individual in respect of all the correlated abilities. The second part has been called the 'specific factor' and denoted by the letter *s*. It not only varies from individual to individual, but even for any one individual from each ability to another. The proof of this all-important mathematical theorem has gradually evolved through successive stages of completeness, and may now be regarded as complete.

"Although, however, both of these factors occur in every ability, they need not be equally influential in all. On the contrary, the very earliest application of this mathematical theorem to psychological correlations showed that there the *g* has a much greater relative influence or 'weight' in some of the abilities tested than in others. Means were even found of measuring this relative weight. At one extreme lay the talent for classics, where the ratio of the influence of *g* to that of *s* was rated to be as much as 15 to 1. At the other extreme was the talent for music, where the ratio was only 1 to 4."

Now this general factor *g* is "not any concrete thing but only a value or magnitude." We have not discovered what it is

like but only where it can be found. "It consists in just that constituent—whatever it may be—which is common to all the abilities inter-connected by the tetrad equation." Spearman finds that the tetrad equation holds good for the inter-correlations of all sorts of abilities. *G* is universal, and hence its great importance. We can now see why a miscellaneous group of tests like the Binet Scale is successful. It gives us a rough measure of *g*. "What the pooling does effect is to make the influences of the many specific factors more or less neutralize each other, so that the eventual result will tend to become an approximate measure of *g* alone."

The tetrad equation does not always hold. Under certain circumstances it ceases to be valid. The most important case is where the *s* factors are very much alike. If two tests resemble each other very materially, such as the test of cancellation of *a's* and the test of cancellation of *e's*, then there is a great deal of overlap between the *s* factors, and the tetrad equation is not fulfilled. This has given rise to the use of the tetrad equation for the discovery of so-called group factors.

"Overlapping specific factors have since often been spoken of as 'group factors.' They may be defined as those which occur in more than one but less than all of any given set of abilities. Thus, they indicate no particular characters in any of the abilities themselves, but only some kinship between those which happen to be taken together in a set. Any element whatever in the specific factor of an ability will be turned into a group factor, if this ability is included in the same set with some other ability which also contains this element. The most that can be said is that some elements have a broader range than others, and therefore are more likely to play the part of group factors."

The proof that *g* exists is based upon a study of many groups of inter-correlations between all sorts of "intelligence" tests. Such tests show a normal distribution of tetrad differences. The tetrad differences arising from the inter-correla-

tions of physical measurements show no such normal distributions and hence cannot be explained by means of the two factors of *g* and *s*.

Spearman then takes up the question as to "how far the range of such tests obeying the tetrad equation really extends." This brings him to "a general survey of the entire range of possible operations of knowing." So he takes up in order a treatment of his three ultimate qualitative laws (Spearman, 23). The first law, the apprehension of one's own experience, has as yet no appropriate measures and hence we know nothing about the presence of *g* in that field. With reference to the other two laws, the eduction of relations and eduction of correlates, we are more fortunate. There are many tests in these two fields of knowledge, although by no means a systematic covering of the fields. "Still, throughout nearly all the immense area into which these tests do appear to have penetrated, the presence of *g* has been decidedly affirmed."

"Does this result, it may be asked, admit of reduction to the simple formula, that *g* measures the power to grasp relations? The answer must be in the negative. In the first place, such a formula would suggest only the educing of relations, and would therefore leave out of account the power—at least equally important—of educing correlates. In the second place, it would overlook the possibility, indeed probability, that *g* also enters into the power indicated by the first neogenetic law, that of knowing one's own experience. Yet again, it would unjustifiably imply that *g* constitutes the whole of any such power, whereas the evidence indicates that *g* is never more than a factor in it."

We find then that *g* exists and that it is well-nigh universal among mental abilities. It is the general factor entering into any performance, whereas *s* is the specific capacity for that particular kind of performance. We do not know, however, the essential nature of what is measured by *g*. Psychologically we may think of it as mental energy.

"In short, although there seem to be grounds for hoping

that a material energy of the kind required by psychologists will some day actually be discovered—whereby physiology will achieve the greatest of all its triumphs—still there is no reason why such energy should have more than a broad analogy to anything of the kind that has been suggested hitherto. Consider, for comparison, how far off the modern concept of electricity is from the old two fluids of Symmer; yet, by virtue of certain analogies, his view is still to this day found sufficiently near the truth to afford the most convenient concept for the purpose of instructing children. For our present purpose, promising ideas are being put forward already, notably by Head and Myers. The latter writes quite definitely: 'I see no reason why we should not identify central nervous energy with mental energy.' "

Physiologically we may think of it as nervous energy, and *g* is due to the fact that each "focus of cortical activity receives continual support from energy liberated by the entire cortex." But we do not know. There is no proof of this. As Spearman says, physiologically *g* may be due to the general plasticity of the nervous system, the quantitative and qualitative alterations of the common blood supply, the state of the endocrine glands or many other physiological conditions.

There is, of course, no need for framing any hypothesis as to the nature of *g*. The proof that *g* exists would still remain, with or without an hypothesis as to its nature. There is, however, according to Spearman, one hypothesis and only one that appears to fit all the facts known at present.

"This is to regard *g* as measuring something analogous to an 'energy'; that is to say, it is some force capable of being transferred from one mental operation to another different one. Even on the physiological side, there are some grounds for hoping that some such energy will sooner or later be discovered in the nervous system, especially the cerebral cortex. . . . But all energy needs to be supplemented by some engine or engines in which to operate. And such engines are obviously supplied by the nervous system, in so far as its func-

tion is localized. Incidentally, this leads to the suggestion that cerebral localization serves three main purposes, sensation, movement, and retention.

"Some of us may be inclined to take yet another step and think that, where energy and engines operate, there must furthermore exist an engineer. And this requirement also seems to be met, namely, in the conative law as expounded in Chapter XIX and as specially manifested in *W*."

2. *Objections to Spearman's Theory.* Objection has been taken to Spearman's theory by several psychologists and particularly so by Thomson (Brown and Thomson, 21). He proposes a sampling theory of ability in which a number of factors are present in the carrying out of any activity such as a mental test, and these factors are a sample of all those which the individual has at his command. This theory makes fewer assumptions than does the special form of theory proposed by Spearman. "It does not," says Thomson, "deny general ability, for if the samples are large there will of course be factors common to all activities. On the other hand it does not assert general ability, for the samples may not be so large as this, and no single factor may occur in every activity. . . . The Sampling Theory, then, neither denies nor asserts General Ability, though it says it is unproven."

Furthermore, Kelley (28) in analyzing the results of many tests given to several grade populations, finds a general factor. But an analysis of this factor "leads us to wonder whether, had we experimentally allowed for maturity, race, sex and general nurture, any general factor would have remained. In other words, we may wonder if there is any factor at all independent of these things corresponding to Spearman's idea of a central fund of intellective energy, or general ability, or *g*." And, finally, Kelley concludes that all of his data, as well as the data of many other workers, point to a multiple-factor hypothesis rather than to a single-factor hypothesis.

3. *Thorndike's Theory.* Thorndike's definition of intelligence and his contribution to a description of intelligence by

the attributes of speed, level, range and area, have already been mentioned.

Thorndike (27) approaches the problem of the nature of intelligence or intellect from the empirical point of view. Intellect is the ability to succeed with certain tasks. We measure it by taking a fair sampling of these tasks. "We contrast intellectual power over things as by ideas about length or weight or heat, with non-intellectual power over things as by strength or skill or acuity of vision. We contrast intellectual power over people, as by consideration of facts about them, with non-intellectual power over them, as by good temper or courage or physical charm. We contrast intellectual power over ideas, as by using other ideas to gain success with them, with non-intellectual power over them, as by industry or patience."

He then comments on the fact that it has been customary to make a rather sharp distinction between mere connection-forming and the association of ideas on the one hand, and the higher thought processes, which include abstraction, generalization, and reasoning, on the other. He then continues, "The hypothesis which we present and shall defend admits the distinction in respect of surface behavior, but asserts that in their deeper nature the higher forms of intellectual operation are identical with mere association or connection forming, depending upon the same sort of physiological connections but requiring many more of them. By the same argument the person whose intellect is greater or higher or better than that of another person differs from him in the last analysis in having, not a new sort of physiological process, but simply a larger number of connections of the ordinary sort."

And further, "Our hypothesis limits itself to the original capacity. If by original nature, apart from all training, a man possesses tendencies to be right rather than wrong in his judgments, to hold true rather than false ideas, to make justifiable rather than unjustifiable inferences, more or less

than other men, in so far forth those tendencies are due to his having more or fewer *c's* than other men.

"The essential element of our hypothesis is that it offers a purely quantitative fact, the number of *c's*, as the cause of qualitative differences either in the kind of operation (e.g., association versus reasoning) or in the quality of the result obtained (e.g., truth versus error, wisdom versus folly), so far as these qualitative differences are caused by original nature."

If, then, one man differs from another with reference to having more intelligence or a greater fund of ideas, what is it due to? And Thorndike answers, "This greater fund of ideas and connections is partly due to larger life and more varied and stimulating life, but it may be and certainly is partly due to original nature. It has some anatomical or physiological cause or parallel. Our hypothesis regards this anatomical cause or correspondent of the original possibility of having more such connections (call it *C*) as the cause of the original differences in intellect among men. . . . What is essential to the hypothesis is that by original nature, men differ in respect of the number of connections or associations with ideas which they can form, so that despite identical outside environments, some of them would have many more than others. 'The number of *c's* a man has' means simply the original constitutional basis of the number of ideational connections which he has. . . . Negatively, the hypothesis asserts that no special qualitative differences are required to account for differences in degree of intellect; the higher processes or powers have no other basis in original nature than that which accounts for differences in the number of bonds of the associative type. . . . The gist of our doctrine is that, by original nature, the intellect capable of the highest reasoning and adaptability differs from the intellect of an imbecile only in the capacity for having more connections of the sort described."

The high correlations between tests are due, then, to the fact that the more difficult items of all types of tests re-

quire more *c's*, and the less difficult fewer *c's*. Perfect correlation does not exist because of differences in interest among different subjects, or because of differences in training or previous background.

There is, thus, in Thorndike's view of the problem no need for two factors, general and specific. Yet at the same time Thorndike's theory of intelligence would not exclude the possibility of a general factor such as Spearman has proposed.

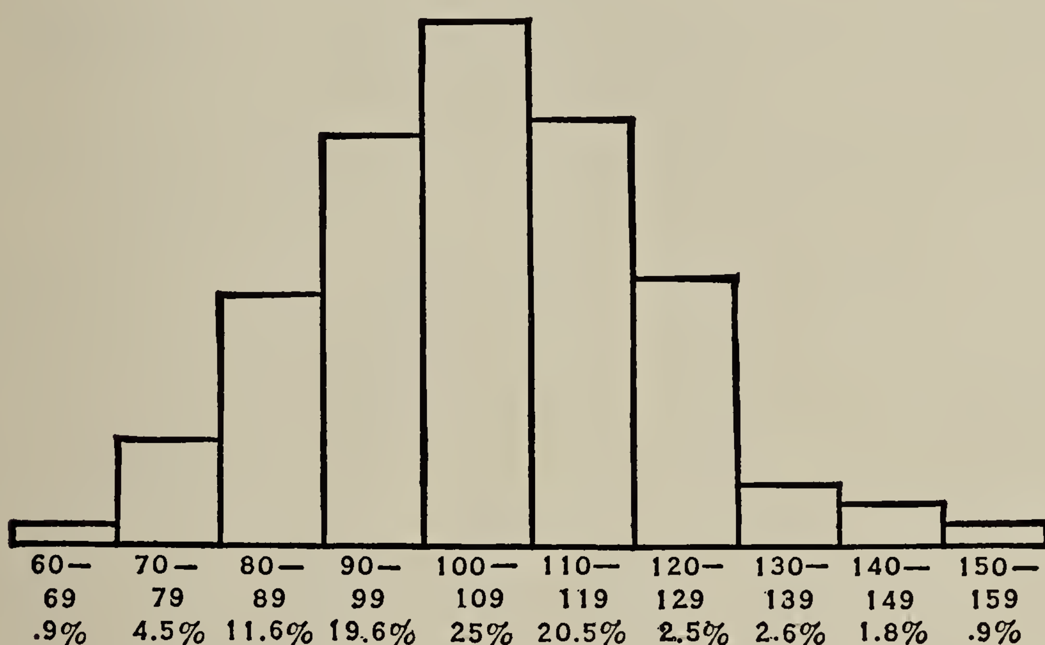


FIG. 2.—The Distribution of I.Q.'s of Kindergarten Children. (From Terman's *The Intelligence of School Children*. Houghton Mifflin Company, publishers.)

If it should eventuate that there is a general factor common to all intellectual processes, then in Thorndike's terminology, this would mean that certain *c's* are common or necessary to all intellectual tasks. Group factors, rather than general factors, would mean that certain combinations of *c's* are necessary for certain tasks and other combinations for other tasks. The Thorndike theory does not regard the Spearman theory as an important issue. Thorndike has concentrated rather upon the task of the measurement of intelligence and a fuller description of intelligence itself, with reference to the attributes of altitude, range, area and speed.

Distribution of Intelligence.—It used to be thought that great intellects, such as geniuses, were widely separated from the mass of mankind. And again, those of low intellect, such as the idiots and imbeciles, were considered a separate species. Average or normal people stood in between these two extreme groups, and they were thought all to have about the same

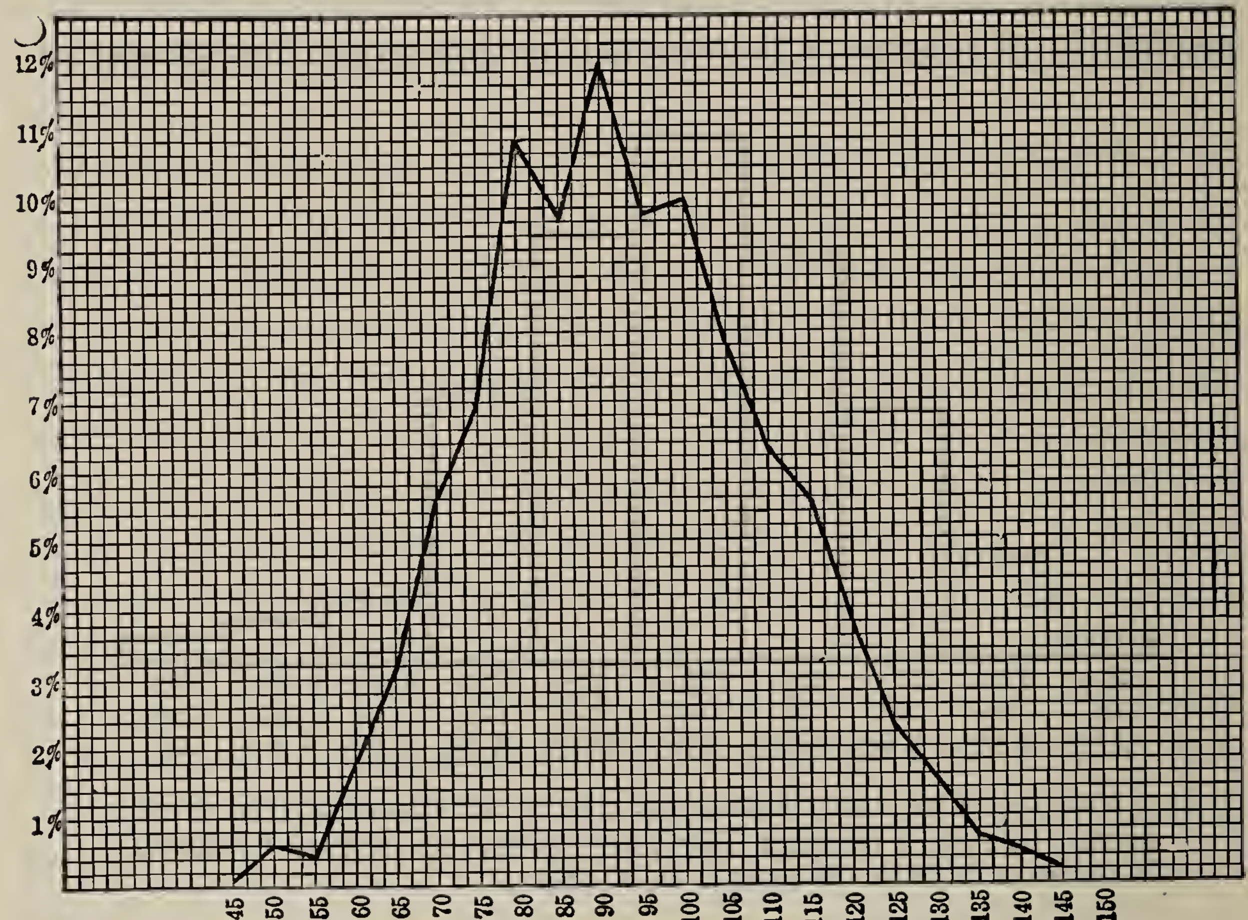


FIG. 3.—Showing distribution of National Intelligence Test I.Q.'s of 1,225 pupils above Grade IV, Public School 64. (From Irwin and Marks' *Fitting the School to the Child.* By permission of the Macmillan Company, publishers.)

amount of intellect, differences in achievement being due to differences in effort or opportunity. So it was considered the proper and desirable thing for the schoolmaster to flog all who did not learn their lessons, that is, all except the unmistakable idiots or imbeciles, who were excused from school and allowed special privileges. This tripartite arrangement of mankind with reference to intellect formed the underlying assumption of much of our educational practice until very recently, and

the influence of this belief is still with us. The results of intelligence testing show a very different arrangement of intellect. Whenever intelligence tests are given to a large un-

Fig. 4.—I.Q. distribution for 999 gifted and 905 unselected children. (From Terman's *Genetic Studies of Genius,* Vol. I. Stanford University Press.)

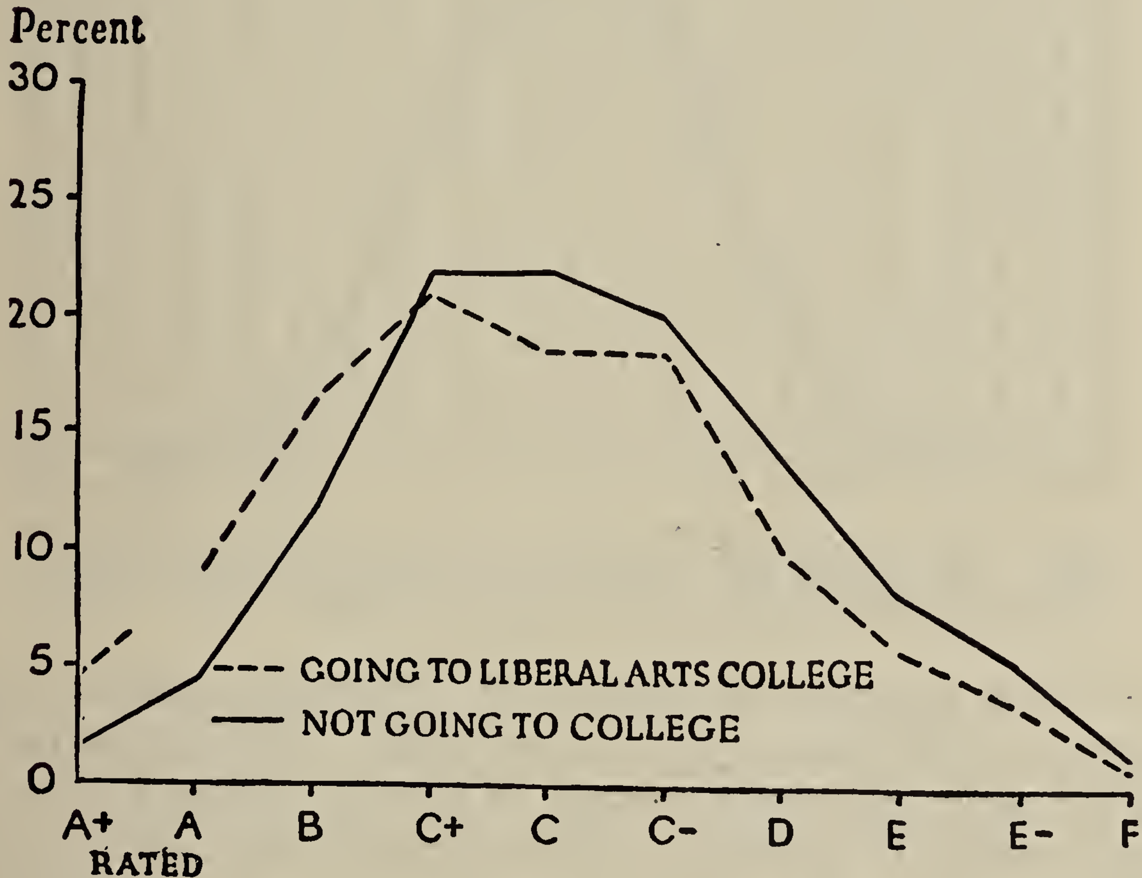


Fig. 5.—Frequency curves for high school seniors. (From Book's *The Intelligence of High School Seniors.* By permission of the Macmillan Company, publishers.)

selected group of individuals and the mental ratings of the group plotted on a frequency curve, we find that these curves tend to take a certain form. If we study Figures 2 to 8, we

note that the general shape of the curve is more or less the same in all of them. The curves begin low at the left, rise gradually higher and higher until they reach their maximum in the middle and then gradually sink again to a low point at the right. In terms of an intelligence test this means that there are a few people who make very low scores and that as

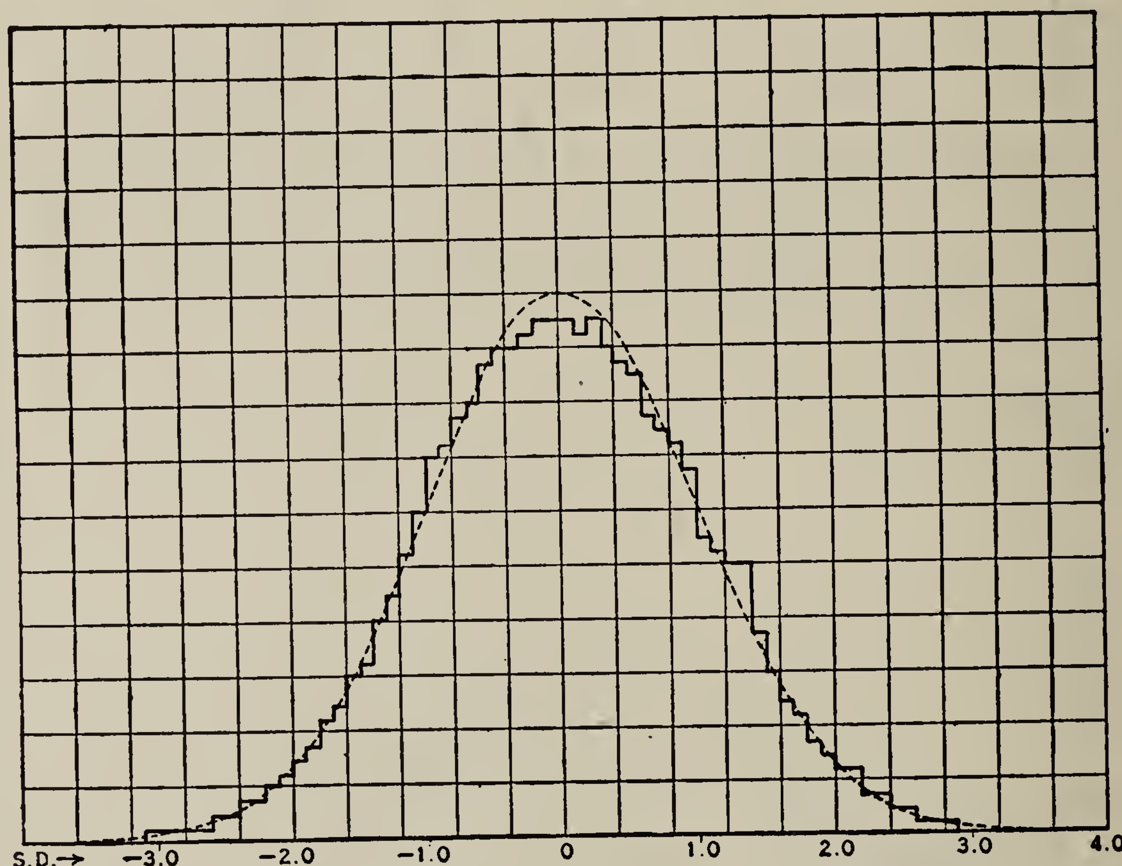


FIG. 6.—Composite curve for the ninth grade based upon eleven tests. The broken line indicates the theoretical normal curve. (From Thorndike's *Measurement of Intelligence.*)

the scores increase the number of individuals making such scores also increases up to a given point. This highest point on the curve represents average or normal ability, because a great number of individuals cluster around this point. From this point on, as the scores increase, the number of individuals becomes less and less until we end up with few individuals making extremely high scores. In other words, the frequency of very low and very high intelligence scores is small. There are few idiots and few geniuses in the world. The frequency of the intermediate levels is somewhat greater. There are

more dullards and bright people. Most numerous of all are the people of average intelligence, who cannot be called either dull or bright and who are far from being idiots or geniuses. Normal intelligence is that degree of ability possessed by the median or average individual, and, because intelligence is distributed as we have described, it follows, therefore, that

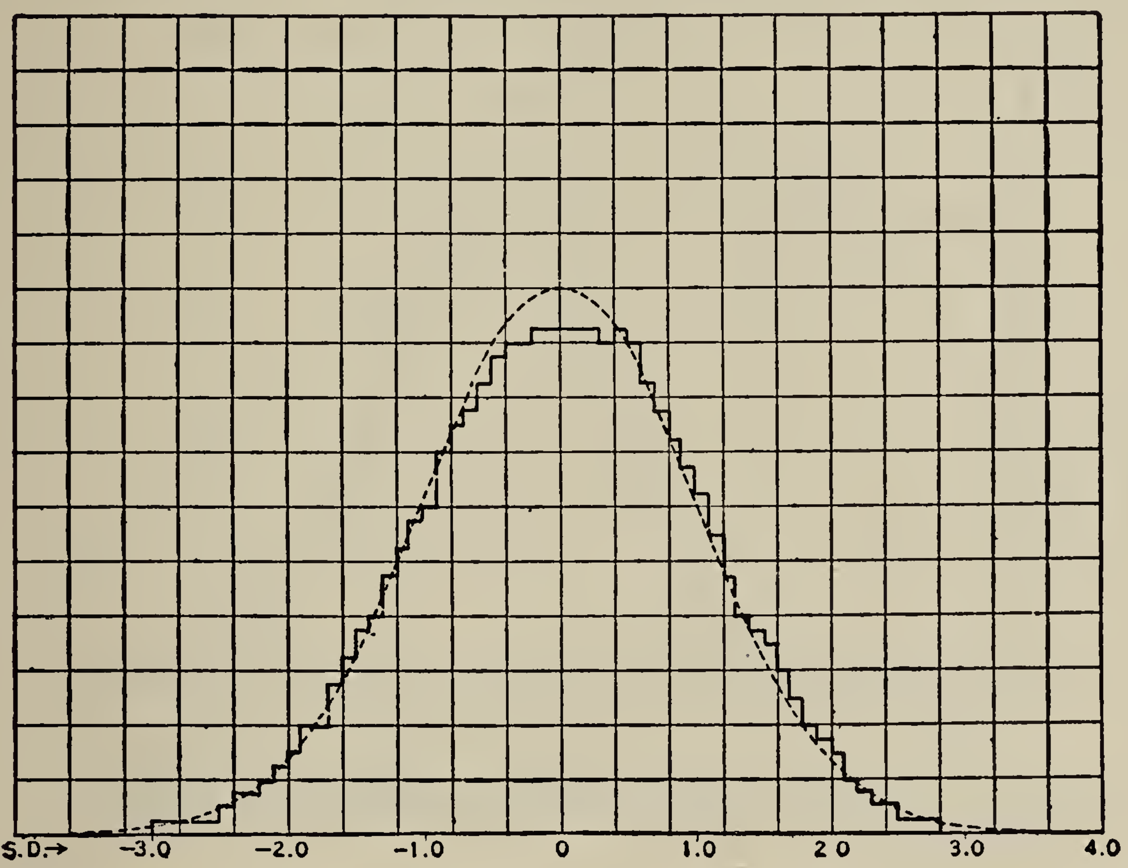


FIG. 7.—Composite curve for college freshmen, derived from eleven single curves. The broken line indicates the theoretical normal curve. (From Thorndike's *Measurement of Intelligence.*)

normal intelligence is the most common or frequent.

The type of distribution we have been describing is generally called a normal distribution and the curves shown in Figures 2 to 8 are generally referred to as normal curves. Whenever a large number of individuals is tested the tendency is to get a normal curve, even though the group is not a random selection of the population in general. For example, the distribution of the intelligence of high school seniors shown in Figure 5 has the characteristics of a normal curve, and it shows that an unselected sampling of high school seniors will

give us a few seniors of poor intelligence for high school seniors, a few of very exceptional intelligence, with the great majority lying in between these two extremes. Figure 4 shows two curves, one for unselected children and the other for gifted children. The latter curve shows how the lower portion is suddenly cut off owing to the fact that no child below an I.Q. of 135 was included in the gifted group.

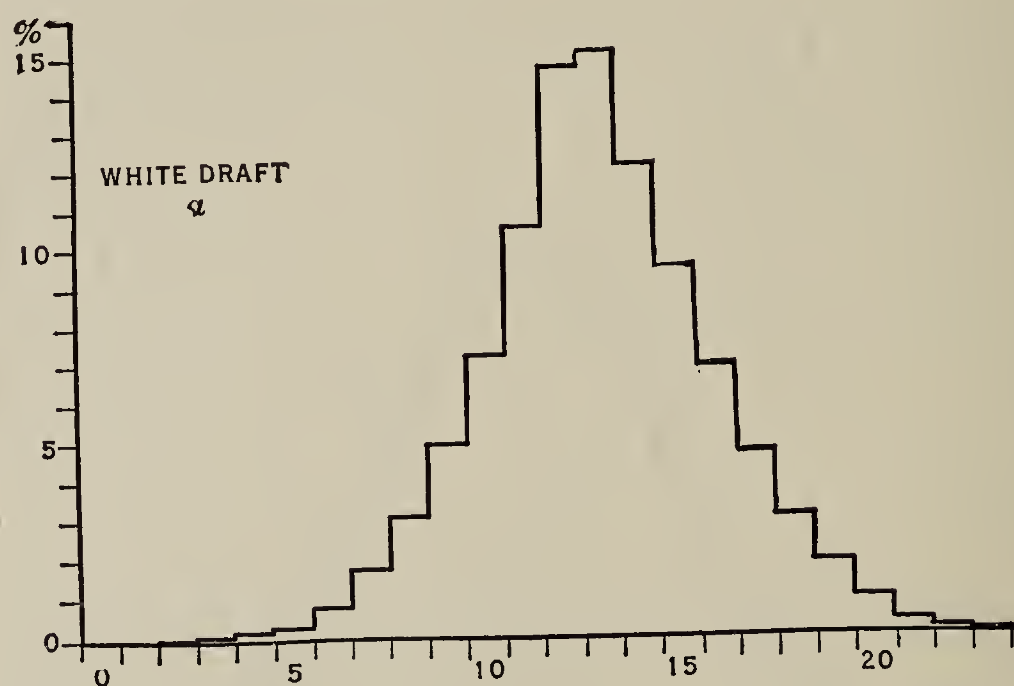


Fig. 8.—Distribution of 93,965 soldiers on combined scale. (From *Memoirs of National Academy of Sciences*, Vol. XV.)

The most thorough study of the distribution of intelligence has been made by Thorndike (27). He presents distributions for single tests for several grade populations and then combines these single curves into composite curves. Two examples of such composite curves are shown in Figures 6 and 7. They conform very closely to the theoretical Gaussian or normal curve.

The Growth of Intelligence.—Everyone knows that as a child grows, he becomes able to do more and more difficult things. Things it was impossible for him to do or understand at three years are readily done or grasped at eight. The capacity for reacting to more and more complex situations

increases as the nervous system develops and grows. At the same time we know that the rate of growth diminishes as the child grows older. This general growth of intelligence is pictured by Thorndike (27) in Figure 9.

This curve shows that the increment of intelligence added each year becomes gradually less and less as the child increases in chronological age. Common observation would seem to bear this out. From birth to one year, the mental change, as well

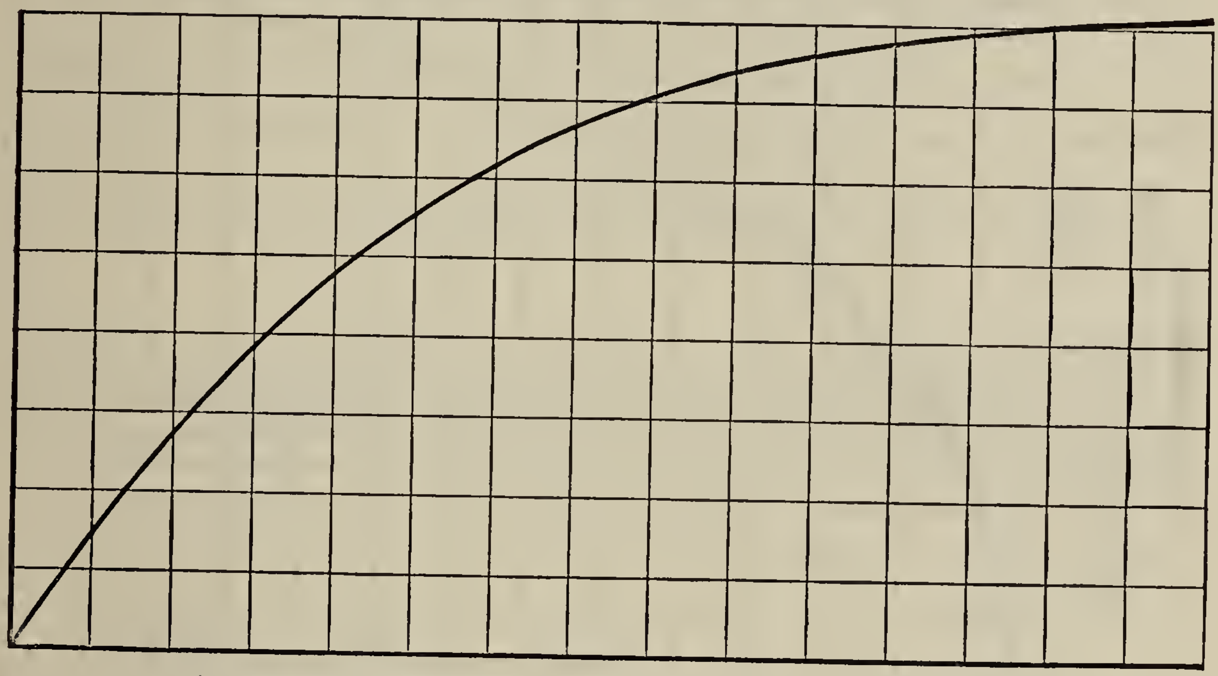


FIG. 9.—The general nature of the relation of altitude of intellect to age in years, 0 to 20. (From Thorndike's *Measurement of Intelligence.*)

as the physical, in an infant is enormous. One can almost literally see him growing and learning to do new things every day. This marvellously rapid rate gradually slows up. We do not note such tremendous changes between ages 7 and 8, or ages 9 and 10. Some have maintained that there is another definite spurt at adolescence, but this has not been clearly demonstrated. Again it is easier to find intellectual tasks which will discriminate between the 5 and 6 year old than to find such as will discriminate between the 12 and 13 year old. The increment of intelligence added each year becomes less and less until it ceases altogether. The curve, as we have said, is a theoretical one. There is no definite proof, because we

lack equal units for the measurement of intelligence. But the assumption of decreasing increments of growth seems to be the most reasonable one.

For the measurement of growth in early infancy we have the notable work of Gesell (28). He made repeated examinations of about one hundred infants. Some children were ex-

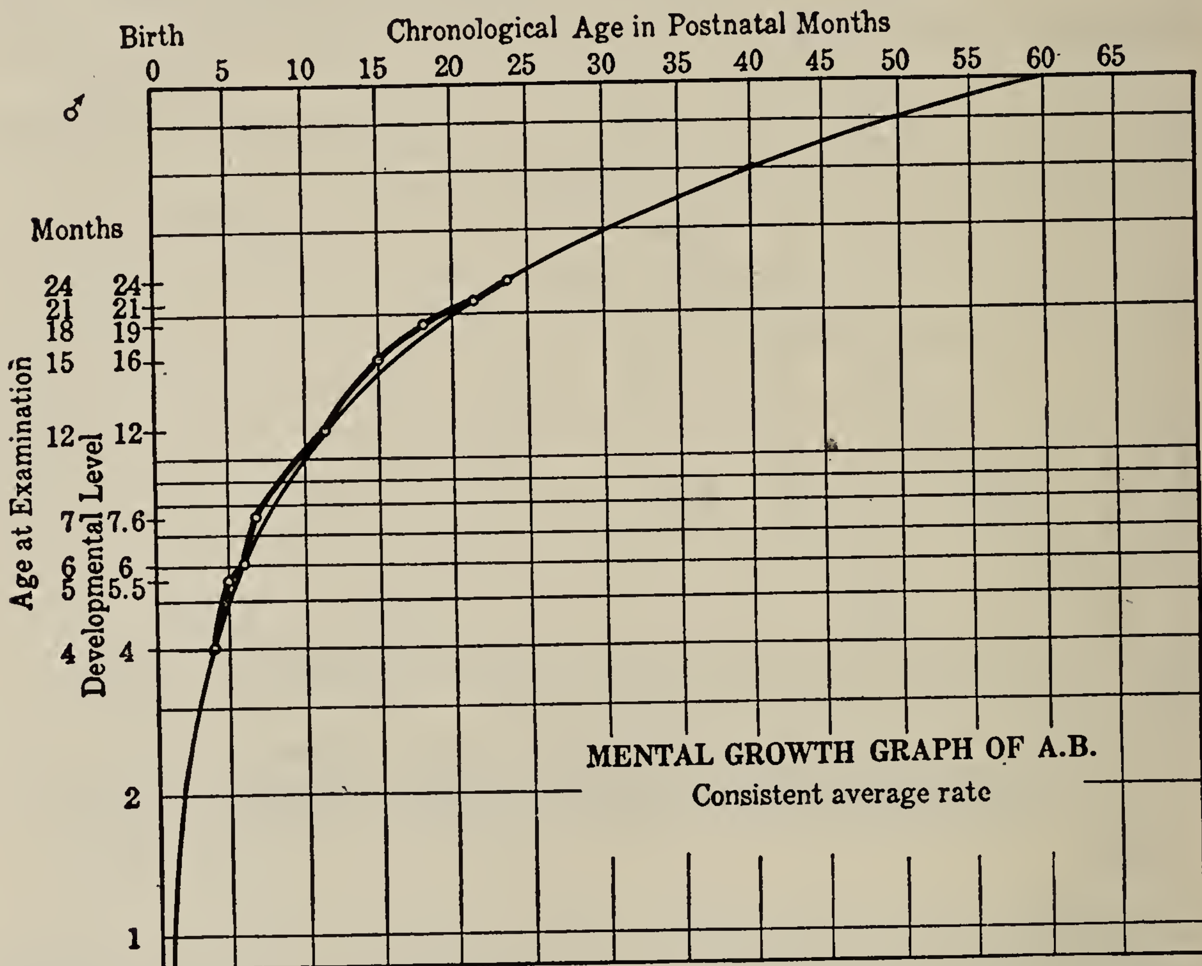


FIG. 10.—Mental growth curve from four to twenty-four months. Infant tested nine times. (From Gesell's *Infancy and Human Growth.*)

amined as often as ten times. His results are expressed in terms of developmental levels and are plotted on a logarithmic chart. Figures 10, 11, and 12 show three of Gesell's curves.* Figure 10 shows the curve of an infant who has been examined nine times. The first column shows when these examinations took place, and the second column shows the developmental

* From Gesell, A. *Infancy and Human Growth.* By permission of the Macmillan Company, publishers.

level attained at each examination. The heavy curve shows the actual measured growth of the infant, and we note that it follows very closely the hypothetical curve of normal growth. Figure 11 shows another case examined six times. The broken line is hypothetical. This case shows consistently subnormal growth. Figure 12 shows a case examined ten times. This

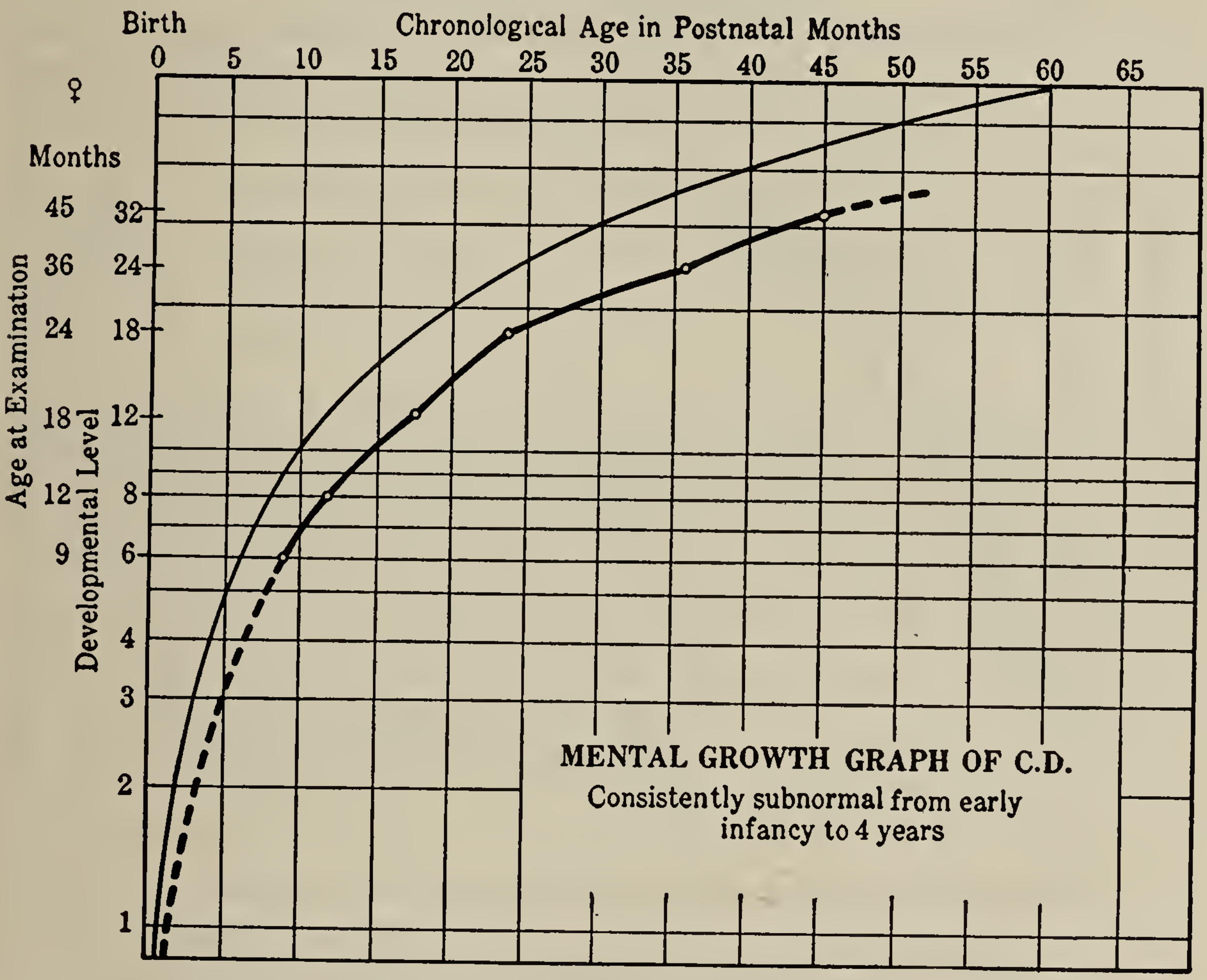


FIG. 11.—Mental growth of sub-normal child from nine months to forty-eight months. (From Gesell's *Infancy and Human Growth.*)

curve remains consistently above the hypothetical normal curve. It shows superior intelligence maintained steadily during the first 24 months of life.

These three curves are typical of the results obtained by Gesell. Consistent, regular growth, whether normal, accelerated or retarded, seems to be the general rule. Nevertheless, there are cases of irregular growth due to all kinds of complications, and Gesell also shows samples of such children.

The best way of measuring mental growth is by repeated measurements of the same child, such as we have discussed above. For older children such repeated measurements are not numerous. Those that are available, are mostly concerned with the Binet test, and if we were to plot them on

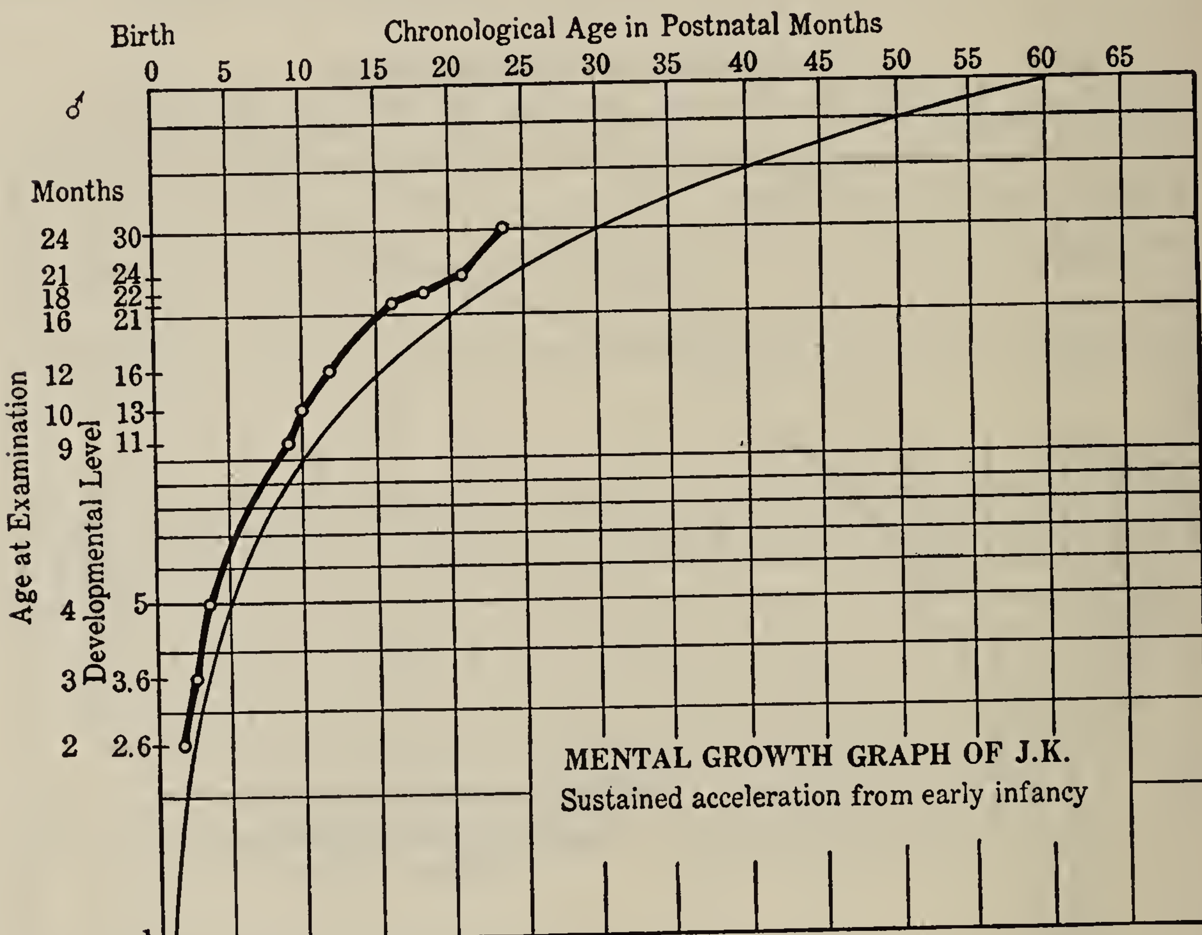


FIG. 12.—Mental growth of superior child from two to twenty-four months. (From Gesell's *Infancy and Human Growth.*)

logarithmic charts, such as Gesell uses, we should obtain similar curved lines showing growth slowing up as the child grows older. The use of the logarithmic charts assumes decreasing acceleration with increasing age.

If then we turn to other tests which give results in terms of score points and not in terms of mental age, what can we learn about mental growth? A study of the average scores by age of many intelligence tests shows in some a tendency for acceleration to decrease slightly between ages 8 and 14;

in others, however, no such tendency is seen, and the curve remains absolutely straight. None show acceleration as age increases. We must note, here, the great difficulty of arriving at any definite conclusion from such a study. If we take the average scores at each age as indicative of mental growth, we make two assumptions: first, that the score units are equal at all levels of the test, which is very doubtful; second, that the selection of children at each age is the same, which is also very doubtful. The real curve of mental growth will not be obtained, until we have a measure of intelligence with equal units at all levels and until we have repeated annual measures of a great number of cases. Until these conditions have been fulfilled the best hypothesis seems to be the one we started with as pictured in Figure 9.

The Limit of Growth.—The growth of intelligence gradually decreases until it reaches a limit. The yearly increments become smaller and smaller, until they become so small as not to be measurable. At about what age does this occur? An exact answer to this question is bound to be very difficult, and is really of no great practical importance. We know the curve gradually decreases in acceleration, and this very gradual decrease would demand the finest instruments of measurement to determine the point at which it stops. We may never know this point, because finer and finer instruments of measurement would continually shift it one way or the other.

Why, then, has this question been so much discussed? The answer is because of the widespread adoption of the I.Q. technique. We cannot go on using the C.A. for the calculation of I.Q.'s beyond the limit of intelligence growth. If we were to do so, I.Q.'s beyond this point would decrease with age and we would have to interpret them differently for each C.A. If we wish to retain the same meaning for a given I.Q. at all ages, it becomes necessary to fix the point at which intelligence growth ceases and then use the same C.A. from that point on for individuals of all ages. It is this practical problem that has pushed the question of the limit of mental

growth into the foreground. Some tentative answer must be given immediately, if we are to continue to use the I.Q. technique for older children. As a matter of fact our problem is split into two, namely: (1) what C.A. must we use for the I.Q. of older children with our present instruments of measurement? (2) At about what age does mental growth seem to reach its apogee? And our answer to the first question would be, use C.A. 14 for most tests as they are constructed and standardized at present; and to the second question, the evidence is not clear; the point may lie anywhere from 14 to 22. Let us see why these two answers seem reasonable.

The I.Q. came into common use through the Stanford Revision of the Binet Scale. In his standardization Terman (17) used elementary school children up to the ages of 13 or thereabout. In order to standardize tests for the higher levels of intelligence high school students had to be used. These tests were assigned to the mental ages of 14, 16 and 18. But obviously the meaning of mental age here is not the same as for the lower ages where the selection of cases is much wider. Then Terman found that his group of "normal" adults, made up of business men and high school pupils, reached on the average an M.A. of 16. He thus recommended the use of 16 for the calculation of I.Q.'s for C.A. 16 and above. This, then, was the first practical answer as to how to handle the I.Q. for older individuals. It was not meant to be a careful scientific attempt to discover the limit of mental growth. Obviously, if the group considered to be normal adults by Terman were not so, there would be an error in the calculation of I.Q.'s.

It would seem from later work, that in all probability Terman's "normal adults" were really above normal in intelligence. The effect of this was to depress the I.Q.'s for adolescents above 14 and for adults. One result of the use of Terman's standard was seen in the early reports of intelligence testing in institutions for delinquents. A disproportionate number of the older as compared with the younger delinquents

were considered feebleminded. Age sixteen is too high an age to use for the Stanford Revision.

Then came the results of the group and individual tests given in the army (Yerkes, 21). The Stanford Binet was given to 653 English speaking men and the mean mental age was found to be 13.42. Evidently the average adult score on the Stanford is more nearly M.A. 14 than M.A. 16. Terman's "normal adults" were probably somewhat above normal.

This same group of 653 men made a mean score of 62.8 on the Army Alpha Group Test. The median score of 51,620 native white soldiers on this same test is 59, so that we may conclude that the small sample of 653 tested on the Binet was probably fairly representative of normal adult intelligence. If we now take this score of 59 or 60 as indicative of average adult performance, we find that it corresponds to what thirteen-year-old children do (Proctor, 21), or it falls between 12 and 13 (Lufkin, 21) or between 11 and 12 (Doll, 19). The median score for 1,108 high eighth grade pupils was found to be 77 (Dickson, 23).

Again we have similar data from Army Examination *a*, a group test used before the Army Alpha was constructed. The median score on this test for 56,140 men was 161. When given to 2,543 school children distributed in age from 7 to 18, we find that a score of 161 falls between the medians for ages 13 and 14. A score of 161 also corresponds to the median score for the seventh grade. This same score falls between mental ages 12 and 13 for a group of 310 school children distributed according to M.A. on the Stanford Binet, but here the number of cases at each age is very small, e.g., 25 at M.A. 12 and 35 at M.A. 13.

From these results it would look as if average adult mentality as represented in the army achieves about as much as children of age 14 or thereabout on our present tests of intelligence. For this reason it would seem to be desirable to use 14 as a basis for the calculation of I.Q.'s for older individuals. Unless we do this we shall find the I.Q. for older groups

declining with our present standardizations of tests. For example if we calculate the I.Q.'s of high school children by using 16 we find that the median I.Q. tends to decrease from the first to the fourth year, a result absolutely contradictory to what we have every right to expect. The duller are eliminated as they go through high school. Their number certainly does not increase. If, on the other hand, we calculate the I.Q. by using 14 we find a tendency for the I.Q.'s to increase, as we should expect. Samples are shown in Table 2. Schools 1 to 4 are very select private schools; school 5 is a public school. School 5 is the only case showing an increase of the median I.Q. by both methods of calculation.

TABLE 2

MEDIAN I.Q.'S OF FOUR YEARS OF HIGH SCHOOL

1. Calculated by using Age 16 as upper limit.

Year	*School* 1	2	3	4	5
I	111	116	117	118	91
II	111	113	113	115	94
III	106	107	111	116	95
IV	109	114	111	112	105

2. Calculated by using Age 14 as upper limit.

Year	1	2	3	4	5
I	113	116	121	121	101
II	117	124	122	123	107
III	120	124	125	130	109
IV	123	130	127	129	120

It would seem, therefore, best at present with our rough tests as they are now standardized to use C.A. 14 as a basis for the calculation of I.Q.'s for all cases above this age.

This does not mean that we have settled the more difficult problem as to just where mental growth ceases. The evi-

dence for any given point is very conflicting. If we take high school pupils in grades 9 to 11 (Thorndike, 23 and 26), we find a measurable increase of about 10 months' mental age by repeating a test after one year interval. But obviously these children are much above average mental ability, and it is likely that children of superior intelligence will tend to grow mentally for a longer period than children of average mentality (Woolley, 26). When all the children of ages 14 and 15 are tested in a city (Hopkins, 24), no increase in M.A. is found from age 14 to 15. When an unselected group of children of all ages is tested (Teagarden, 24), growth would seem to continue up to age 18. Many tests show little or no increase after 15 or 16 (Ballard, 21, Whipple, 24). Recently a very thorough attempt to trace the curve of mental growth has been made by Jones (see Thorndike, 28), who tested all the individuals of a certain community by means of the Army Alpha. He tested 1,131 individuals between the ages of 11 and 54. The median scores for various age groups are as follows:

Age	11	12	13	14	15	16	17	18	19-21
Median	47	52	59	80	86	92	94	93	94

Age	22-24	25-29	30-34	35-39	40-44	45-49	50-54
Median	90	86	83	84	91	73	72

Professor Jones writes:* "The total score goes up fairly steadily to 16.5, shows very minor gains to 20.5, and by 23 has fallen back below the 16 year level." If we interpret the increase up to 20 as a real increase in intelligence, then the decrease after the age of 20 must be explained in terms of an inferior selection of individuals in the 20's due to migration, or to poorer motivation and similar factors among the older people. This seems reasonable, and these results would seem to agree very well with Thorndike's (28) results as to learning ability, which he finds tends to increase gradually up to the age of 22 or thereabout.

* Private letter to the author.

It would seem, therefore, that the curve of growth of intelligence might be somewhat as depicted in Figure 13. This is as much as we know at present and neither the peak of the curve nor the decrease is at all certain. The most reliable part of the curve is that showing rapid growth from birth to age 15 or thereabout.

The Regularity of the Growth.—In the preceding section we discussed the growth of intelligence for individuals in general.

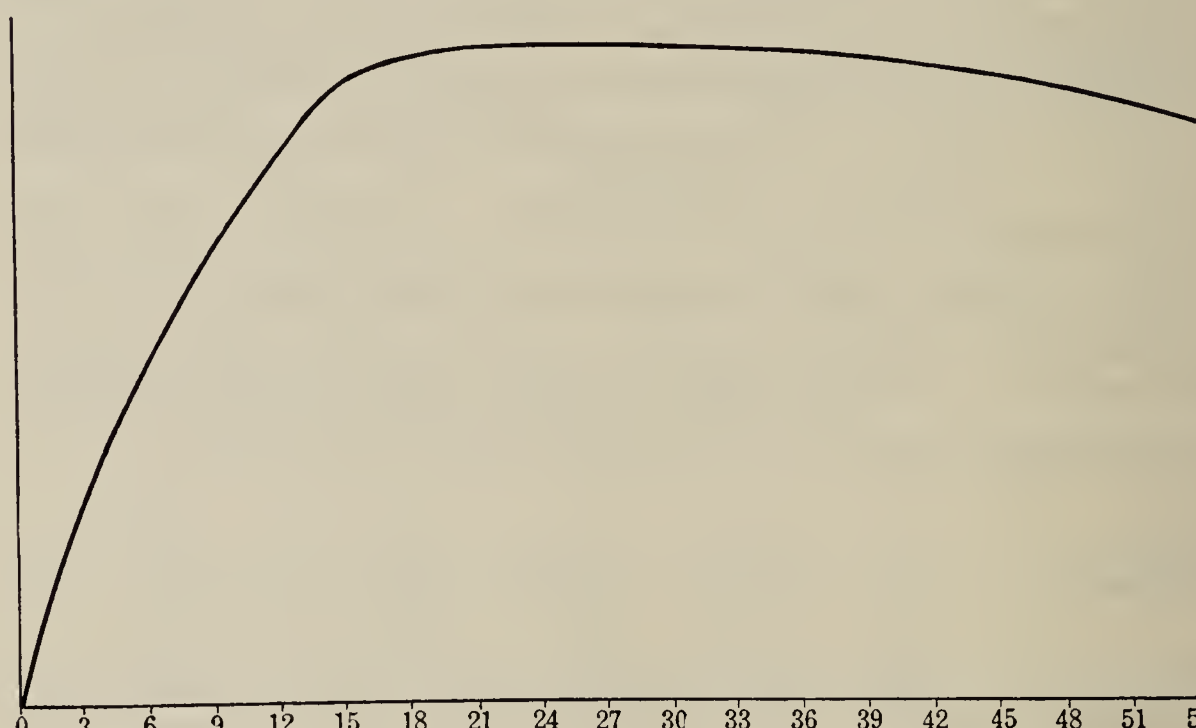


Fig. 13.—Hypothetical Curve of Growth of Average Intelligence.

All our curves were based upon average scores. This means that in general the growth of intelligence is regular from year to year, but it does not mean that every individual's curve will be smooth or regular. Obviously the curves for some individuals might fluctuate. These fluctuations would disappear in the curves based upon averages.

This problem of the regularity of growth is the same as that of the constancy of the I.Q. If the I.Q. is constant from year to year then the curve of growth of intelligence will be regular, because the I.Q. indicates the percentage of growth attained. Let us first examine the results of children tested

several times by the Binet-Simon Tests. Here are the results of a child tested nine times at yearly intervals:

Test	*C.A.*	*M.A.*	*I.Q.*
1	2-3	3-6	155
2	3-1	4-6	146
3	4-0	5-8	142
4	5-2	7-4	142
5	6-0	8-6	142
6	7-2	9-11	138
7	8-2	11-2	137
8	8-11	13-0	146
9	10-0	14-3	143

We notice that the I.Q.'s fluctuate from 137 to 155. The high I.Q. of 155 was obtained when the child was 2 years old, and is based upon very few tests, because the Stanford Binet has no tests below age three and the child's mental age is only 3-6. In all probability, therefore, this examination is a poor one. Neglecting this examination, we note that the other eight I.Q.'s range from 137 to 146. This amount of variation over a seven-year period is relatively small. Such a series of I.Q.'s is called constant. A fluctuating series would show I.Q.'s jumping around from 140 to 100, back to 145 and then down to 90 and so on. Or else we might have a series of I.Q.'s gradually increasing or decreasing from year to year and there are a few cases of such series. In general, however, this is not the case.

Other examples of the re-tests given by Terman (19) show series of I.Q.'s as follows:

Child 1	I.Q.'s	145	150	153
Child 2	I.Q.'s	138	136	144
Child 3	I.Q.'s	144	142	140
Child 4	I.Q.'s	134	141	137

Gray and Marsden (26) give the following examples of I.Q.'s that tend to increase or decrease or are irregular:

I.Q. ON REPEATED TESTS

Child	*1*	*2*	*3*	*4*	*5*	*6*
A	80	89	90	86	79	74
B	101	93	88	88	81	80
C	134	134	132	120	116	117
D	100	104	100	107	111	117
E	104	110	110	97	102	99
F	97	104	98	97	90	93
G	103	108	95	92	106	102

Adams (26-27) reports the results of a case tested twenty-one times. The child was of inferior intelligence and shows a constantly decreasing I.Q.:

Date of Test	*C.A.*	*M.A.*	*I.Q.*
1/13/11	7-3	6-3	86
7/12/11	7-8	6-2	76
10/27/11	8-1	7-0	86
1/ 1/12	8-3	7-0	84
3/ 6/12	8-5	7-1	84
7/22/12	8-7	7-0	80
4/28/13	9-6	6-3	66
6/30/13	9-7	7-2	74
8/ 4/13	9-9	7-2	73
7/17/14	10-8	7-2	67
2/10/15	11-4	7-2	63
1/25/16	12-3	7-2	59
7/26/16	12-8	7-3	57
1/26/17	13-3	8-2	62
7/28/17	13-8	7-3	53
8/14/17	13-9	8-2	59
7/26/18	14-8	8-0	54
8/ 8/18	14-9	8-2	55
11/ 2/23	20-2	8-7	54
11/10/24	21-2	9-0	56
3/24/26	22-5	8-7	54

The amount of constancy or agreement from one test to another can best be expressed by the coefficient of correlation

between the I.Q.'s obtained for the same child by repeated tests. Many workers have reported such coefficients for the Binet Tests. Table 3 shows these results. The first column gives the author's name; the second the date of publication; the third gives the coefficient of correlation; the fourth the number of pairs of I.Q.'s entering into the correlation. In most cases the number of pairs is based upon children tested twice, but in some cases, where children have been tested more than twice, the number is based upon examinations and not children. We note from the table that the coefficients range

TABLE 3

THE CONSTANCY OF THE I.Q. ON THE BINET-SIMON TESTS

Author	*Date*	*r*	*n*	*Comments*
Cuneo and Terman..	1918	.91	77	Av. of 3 r's.
Terman	1919	.93	435	
Rugg and Colloton ..	1921	.84	137	
Garrison	1922	.87	468	Av. of 3 r's.
Gordon	1922	.84	44	
Henmon and Burns..	1923	.91	59	Borderline cases
Berry	1923	.74	351	Retarded children
Gray and Marsden..	1923	.88	218	
Baldwin	1923	.85	143	
Dickson	1923	.90	288	
Madsen	1924	.85	34	
Gray and Marsden..	1924	.85	371	
Rugg	1925	.95	114	
Garrison and Robinson	1925	.90	131	Av. of 3 r's.
Hildreth	1926	.81	1112	
Wentworth	1926	.82	145	
Gray and Marsden..	1926	.85	616	
Randall	1927	.79	152	
Lincoln	1927	.95	30	Two tests on same day
Cushman	1927	.74	144	
Rogers	1928	.72	60	Student testing. Av. of 2 r's.
Brown	1930	.88	707	Problem children

from .72 to .95, the median being .85. The lowest coefficient, .72, is for inexperienced examiners. In most cases the interval between tests is about a year, but one author reports results of two tests given on the same day, and others report re-tests over longer periods, showing no substantial difference in the coefficients. Several authors have also reported median or average differences in I.Q. from test to re-test. This difference is about 5 points.

Similar results for group tests have been reported. Samples of such correlations are shown in Table 4. They are very similar to those in Table 3 for the Binet Scale. The study of Keys (28) seems to indicate that the correlation tends to drop as the interval between tests increases.

TABLE 4

THE CONSTANCY OF THE I.Q. ON GROUP TESTS

Author	*Date*	*r*	*n*	*Interval*	*Test*
Garrison and Robinson . . .	1925	.90	131	10 mos.	National Intelligence
		.91	131	20 mos.	National Intelligence
Bowie and Laws . . .	1925	.87	?	6 mos.	Northumberland
Pintner . . .	1925	.72	26	4 yrs.	Pintner Non-Language
Nettles. . . .	1926	.85	130	3 yrs.	Terman Group
Shewrman.	1926	.77	229	3½ yrs.	Terman Group
Broom. . . .	1927	.93	102	1–23 mos.	Terman Group
Cowdery . .	1928	.75	207	1 yr.	Thorndike Intelligence
Keys.	1928	.81	200	1.3 yrs.	National Intelligence
		.78	200	2.8 yrs.	National Intelligence
		.75	200	4.0 yrs.	National Intelligence
Lamson. . . .	1930	.83	53	5 yrs.	Army Alpha.

All of the results given above indicate that the growth of intelligence is regular rather than fluctuating, but at the same

time a correlation of .85 or .90 leaves plenty of room for individual shifts from one testing to another. What are some of the possible causes for such changes in I.Q.? We may roughly divide them into intrinsic and extrinsic factors, considering under intrinsic, changes in the child himself and under extrinsic those factors concerned with the test itself and the environment of the child.

Intrinsic Factors.—There is no reason why there should not be some individual variations in the regularity of the growth of intelligence, just as there seem to be in the regularity of physical growth. Gesell (28) has presented several cases of infants showing atypical growth. In some cases growth seems to slow up very decidedly, resulting shortly in mental deficiency; in other cases growth seems to increase or take a sudden spurt. Most of these cases are connected with marked physical disturbances. All these cases of Gesell are for very young children. Dearborn (28) gives a few samples of older children showing decided changes in I.Q. He is of the opinion that rate of mental growth is somewhat conditioned by rate of physical growth. De Weerdt (28) cites a case of malnutrition showing improvement in 1½ years from an I.Q. of 116 to 130.

Other workers, however, find no change in the intelligence of school children due to operations for adenoids and diseased tonsils (Rogers, 22), or to the treatment of common minor defects revealed by school medical examinations (Westenberger, 27), or to glandular therapy (Fox, 28). We may sum up, therefore, by saying that some changes in the constancy of the I.Q. may be caused by intrinsic growth factors in the child himself, but that the treatment of most physical defects in the school child seems to lead to no measurable improvement in I.Q.

Extrinsic Factors.—Under this heading we shall take up the following factors which may lead to changes in I.Q. from one test to the next: (1) the coarseness of the scale; (2) the examiner; (3) practice; (4) coaching; (5) environment in general.

(1) *Coarseness of the Scale.* By this is meant that increase in mental age on the Binet Scale takes place in general by jumps of two months. If a child passes an additional test two months are added to his mental age. A few half scores give increments of one month and in the upper ages above age 10, the increments added by passing an additional test are 3, 4, 5, or 6 months. At the lower end of the scale an increment of two months means an increase of 3 or 4 points in I.Q. About year 10 it means an increase of 1 or 2 points in I.Q. At year 14 the passing of an additional test scoring 6 months raises the I.Q. 4 points. It may readily happen, therefore, that on the first examination a child may just pass a hard test and on the second examination just fail a similarly hard test. If this happens about year 10 it will lead to a difference in I.Q. of 4 points; at year 14 it may mean a difference of 8 points. By the very nature of the scale itself, therefore, we are to expect a certain amount of fluctuation from test to re-test.

(2) *The Examiner.* The Binet Scale is not entirely objective. The personality and skill of the examiner play some part in eliciting responses from the child. And, furthermore, in the scoring of several of the tests there may arise legitimate differences of opinion between experts. I.Q.'s are more likely to be constant if test and re-test are given by the same examiner. Hildreth (26) finds a correlation of .87 for tests given by the same examiners, but this drops to .79 for tests of the same child given by different examiners. All of these examiners were more or less expert examiners. We do not include errors in testing or scoring caused by inexpert examiners. Such errors may be very large and may lead to enormous differences in I.Q. Beginners scoring the same test arrive at I.Q.'s differing by 30 points. But these sources of error can be removed by training. There still remains the personal equation of the examiner and the legitimate differences in opinion due to the subjective nature of much of the scoring. Group intelligence tests are generally much less likely to be influenced by the personal equation of the examiner.

(3) *Practice.* Practice on intelligence tests leads to improvement. It would contradict all our laws of learning if it did not. There is nothing magical about the material used in intelligence testing to make it impervious to improvement by practice. If we give the same test or alternative forms of the same test to individuals several times in succession, we notice a tendency for the scores to increase. Several workers have shown that a repetition of Army Alpha may increase the initial score from 8 to 16 points. Thorndike (22) found that college students gained 8 points from trial one to trial two on the Thorndike Mental Alertness Test. By giving fifteen forms, one each day, to school children, he found a general tendency for practice effects to be marked up to the third trial; from then on to the fifteenth the practice effects were very slight.

We have no definite results as to the gains from practice due to taking the Stanford-Binet Test several times on successive days. In all probability there would be a slight gain comparable to that reported for group tests. From our study of re-tests of the Stanford in connection with the problem of the constancy of the I.Q., we know that tests taken a year apart show no effect of practice. Therefore, in the usual school situation, where Binet tests are not given more than once each year and where group tests are never repeated day after day, we may conclude that increases in score due to practice are negligible.

(4) *Coaching.* Somewhat different from the mere repetition of a test, is the deliberate teaching of the items contained in a test or of similar material. Graves (24) has made a thorough study of the effect of coaching on the Stanford-Binet. She finds a very decided increase after definite coaching on the tests themselves or after training on material similar to the tests. About 23 months of mental age can be added by coaching and about 8 by training on similar material. After twelve months the coached group is still slightly superior to the control group. Even after three years some slight effect of coaching may be

present (Greene, 28), but the number of children was then so small as to make this result doubtful.

Many studies of the effects of coaching on group intelligence tests have been made and these have been summarized by Chen (28). All of them show increases in score due to coaching, the amount of increase depending on the intensity of the coaching and also upon the nature of the material. Some test material is more sensitive to the influence of coaching than other material, as Chen shows.

All of this means, from the point of view of intelligence testing, that we cannot interpret in the usual way the scores of children who have been coached. Such children cannot legitimately be compared with the norms, that is, with the children who have not been coached, and, therefore, M.A.'s and I.Q.'s calculated for coached children are meaningless. Or, in other words, the background of the coached group differs so much from the background of the uncoached or standardization group, as to make the two groups not comparable. But the coached children may be compared among themselves, since initial and final scores of coached groups usually correlate fairly high.

In the practical work of intelligence testing in the schools, coaching can frequently be detected. A skillful individual examiner can generally find out whether a child has been taught some of the tests. Although some parents and teachers may coach their children, they rarely take the child into their confidence and ask him to cooperate in deceiving the psychological examiner. The coaching of an individual child in a group test is less easily discovered, unless it be so thorough as to place him head and shoulders above the group. The coaching of a whole class may make it stand out so prominently above other similar classes as to arouse immediate suspicion.

(5) *Environment.* By environmental influences we may mean very specific and narrow changes or else those wider factors of home and school. An example of a very specific and narrow change is seen in the work of Hurlock (25), who found

that praising or censuring children caused a slight increase in score on a group intelligence test as contrasted with the scores of a control group that was neither praised nor blamed. The permanence of such changes in score has not been studied. It is probably slight and will not affect the general intelligence of the child.

With reference to the wider environment of home life and school life, we may say that such shifts as normally occur cannot be very influential, because of the fact that the average child grows constantly and regularly in general intelligence. The differences in the amount of school attendance shown by children in an ordinary city school are not great enough to have any influence on their intelligence ratings (Denworth, 28). But if these differences in the amount of school attendance become very great, then we find a significant drop in I.Q. on the Binet Test. Gordon (23) found that with Gipsy children who attend school only 35 per cent of the time, there is a decided tendency for the I.Q. to drop as the child grows older; and with Canal Boat children who attend school only 5 per cent of the time, this tendency is more marked than with Gipsy children. These results were found with the Binet Scale. If we hold this scale to be a true measure of intelligence at all times for all children, then we must conclude that not going to school regularly causes a great decrease in intelligence. If we consider the Binet Scale merely as a differential measure of intelligence for children of similar background, we must conclude that the background of Gipsy and Canal Boat children is so different from that of the standardization group of the Binet Scale as to make this scale useless for differential purposes in their case. Gaw (24) gave both performance and Binet tests to Canal Boat children in England and found their average I.Q. on the Binet to be 69 as compared with an average I.Q. of 82 on a performance scale. Obviously the latter is a more accurate measure of their intelligence than is the former.

The most noteworthy attempts to measure the general influence of home environment upon the Binet Scale have been

made by Freeman (28) and by Burks (28). Freeman found that the average I.Q. of 74 children before adoption was 91.2, whereas, after they had been in a foster home on the average for four years, this I.Q. increased to 93.7, that is a gain of 2.5 points, or an estimated gain of 7.5 points, allowing for the general tendency of the I.Q. to decrease with age. When these cases were divided into better and poorer homes, he found the increase for the better homes to be 5.3 or, with the allowance, 10.4 points in I.Q., whereas the poorer homes showed an increase of only 0.1, or with the allowance, 5.0 points in I.Q. Another way of showing this increase is by means of the correlation between the intelligence and the home rating, which rose from .34 at the time of placement to .52 four years after. Again, Freeman finds that the correlations between the I.Q.'s of 125 pairs of siblings placed in different homes for a mean period of seven years range from .25 to .34, which is lower than the correlation of about .50 usually found for siblings in the same home.

Similar testimony as to influence of home environment on the Binet rating comes from the work of Burks. She compared a group of 214 foster children placed before the age of 12 months, who were between age 5 and 14 at the time of the investigation, with a control group of 105 true children, the families being matched for home ratings, education of parents, etc. The mean I.Q. of the foster children is 107, as contrasted with an I.Q. for the controls of 115. Burks' estimate of gain in I.Q. points due to environmental shifts is from 3 to 9, as contrasted with Freeman's estimate of from 7 to 10. The total possible increase, according to Burks, from a very poor to an exceptionally good home might be from 9 to 27 I.Q. points.

Although these two investigations do not agree in every detail, they are sufficiently alike to make the point that a decided shift in environmental conditions made at an early age will probably affect the score on the Binet Test. The amount of change made in the score on the Binet Test will depend upon

the age of the child when the shift is made and also upon the extent of the environmental shift. Slight changes in environment will have no effect. The change must be a very decided one, and the child must be young enough to profit by the change.

Both of these investigations show the changes due to marked environmental shifts on the Binet score. If we regard this mental rating as a true measure of innate intelligence under all conditions, then we must interpret the results as showing the influence of environment upon innate general intelligence. If, however, we consider the Binet scale or any other scale merely as a measure of differences in intelligence among children having the same general background or environmental experience, then we must conclude that it fails to measure native ability accurately where environmental backgrounds differ as much as they did in the experiments described above. Fortunately, in Freeman's work, he gave the International Group Mental Test to 374 cases. The difference between the children in the poorer and better homes did not seem to be as great on this test as on the Binet. Because there are no age norms for the International Group Mental Test, a real comparison is difficult, but the suggestion seems to be that a non-language test may be more adequate to measure differences of intelligence between children of very different environments. The Binet may contain too many tests favoring those in good environments. At present, therefore, we may say, either that intelligence is somewhat susceptible to the environmental influences of home conditions, or that the Binet Scale is not a very accurate measure of intelligence for children differing greatly in home conditions. It probably penalizes those brought up in very wretched homes. This latter view-point would lead to the construction of intelligence tests less dependent upon environmental factors, although we shall never be able to make one absolutely independent of environmental factors.

REFERENCES

Adams, M. E. (26-27). "Promise and Fulfillment." *Training Sch. Bull.,* 23, 209-216.

Avery, G. T. (23). "A Study of the Binet and Terman Intelligence Tests with eleven-year-old Children." *J. of Ed. Res.,* 7, 429-433.

Ballard, P. B. (21-22). "The Limit of the Growth of Intelligence." *Brit. J. of Psych.,* 12, 125-141.

Ballard, P. B. (22). *Group Tests of Intelligence.* London.

Baldwin, B. T. (23). "Additional Data from Consecutive Stanford-Binet Tests." *J. of Ed. Res.* 8, 375.

Bell, J. C. and Suhrie, A. L. (28). *Contributions to Education,* Vol. II (N. Y. Soc. for Exp. Study of Educ.). Yonkers, pp. xi + 425.

Berry, C. S. (23). "The Intelligence Quotients of Mentally Retarded Children." *Sch. and Soc.,* 17, 723-729.

Binet, A. and Simon, T. (16). *The Development of Intelligence in Children.* Trs. by E. S. Kite. Vineland Training School.

Bowie, S. and Laws, A. R. (25). "A School Intelligence Test." *Forum of Education,* 3, 32-39.

Bronner, A. F. et al. (27). *A Manual of Individual Mental Tests and Testing.* Boston, pp. x + 287.

Brooks, F. D. (27). "The Accuracy of Intelligence Quotients from Pairs of Group Tests in the Junior High School." *J. of Ed. Psych.,* 18, 173-186.

Broom, M. E. (27). "A Study of the Constancy of the I.Q." *Educ. Res. Bull.,* Los Angeles City Schools, Vol. IX, No. II.

Broom, M. E. (28). "A Note on the Validity of a Test of Social Intelligence." *J. Appl. Psych.,* 12, 426-428.

Brown, A. W. (26). *The Unevenness of the Abilities of Dull and of Bright Children.* T. C. Contribs. to Educ., No. 220, T. C. Columbia University.

Brown, A. W. (30). "The Change in Intelligence Quotients in Behavior Problem Children." *J. Educ. Psychol.,* 21, 341-350.

Brown, W. and Thomson, G. H. (21). *The Essentials of Mental Measurement,* Cambridge Univ. Press.

Burks, B. S. (28.) "The Relative Influence of Nature and Nurture upon Mental Development." *Twenty-seventh Yearbook N. S. S. E.,* Part I, Chapter X, pp. 219-316.

Chen, H. S. (28). *The Comparative Coachability of Certain Types of Intelligence Tests.* T. C. Contribs. to Educ., No. 338.

Cowdery, K. M. (28). "Repeated Thorndike Intelligence Examinations." *Sch. and Soc.,* 27, 367-369.

Cuneo, I. and Terman, L. M. (18). "Stanford Binet Tests of 112 Kindergarten Children and 77 Repeated Tests." *Ped. Sem.,* 25, 414-428.

Cushman, C. L. (27). "A Study of the Reliability of Mental Tests as Used in Oklahoma City." *J. Appl. Psych.,* 11, 509-511.

Dearborn, W. F. (28). *Intelligence Tests,* Boston.

Denworth, K. M. (28). "The Effect of the Length of School Attendance Upon Mental and Educational Ages." *Twenty-seventh Yearbook N. S. S. of E.*

De Weerdt, E. H. (28). "The Effect of Malnutrition Upon Expression of Intelligence." *Sch. and Soc.*, 27, 540-541.

Dickson, V. E. (23). *Mental Tests and the Classroom Teacher.* Yonkers, N. Y.

Doll, E. A. (19). "The Average Mental Age of Adults." *J. Appl. Psych.*, 3, 317.

Dougherty, M. L. (28). *A Comparative Study of Nine Group Tests of Intelligence for Primary Grades.* Baltimore, Johns Hopkins Press.

Edwards, A. S. (28). "Intelligence as the Capacity for Variability or Versatility of Response." *Psych. Review*, 35, 198-210.

Flemming, C. W. (25). *A Detailed Analysis of Achievement in the High School.* T. C. Contribs. to Educ., No. 196, New York.

Fox, E. J. (28). "An Investigation of the Effect of Glandular Therapy on the Intelligence Quotient." *Mental Hygiene*, 12, 90-102.

Freeman, F. N., et al. (28). "The Influence of Environment on the Intelligence, School Achievement and Conduct of Foster Children." Chapter IX. *Twenty-seventh Yearbook N. S. S. E.*, Part I, pp. 101-217.

Freeman, F. N. (26). *Mental Tests.* Boston.

Garrett, H. E. and Kellogg, W. N. (28). "The Relation of Physical Constitution to General Intelligence," etc. *J. of Exp. Psych.*, Vol. XI, 113-129.

Garrison, S. C. (22). "Additional Retests by means of the Stanford Revision of the Binet-Simon Tests." *J. of Ed. Psych.*, 13, 307-312.

Garrison, S. C. and Robinson, M. S. (25). "A Study of Retests." *J. of Ed. Res.*, 11, 190-196.

Gaw, F. (24-5). "A Study of Performance Tests." *Brit. J. of Psych.*, 15, 374-392.

Gaw, F. (25). *Performance Tests of Intelligence.* Report No. 31, Industrial Fatigue Research Board. His Majesty's Stationery Office, London.

Gesell, A. (28). *Infancy and Human Growth.* New York.

Goodenough, F. L. (26). *Measurement of Intelligence by Drawings.* Yonkers.

Gordon, H. (23). *Mental and Scholastic Tests among Retarded Children.* Educational Pamphlets. No. 44, Board of Education, London.

Gordon, K. (22). "Some Re-tests with the Stanford-Binet Scale." *J. of Ed. Psych.*, 13, 363-365.

Graves, K. (24). *The Influence of Specialized Training on Tests of General Intelligence.* T. C. Contribs. to Educ., No. 143, Teachers College, Columbia University.

Gray, P. L. and Marsden, R. E. (23). "The Stanford-Binet Tests in Some English Schools." *J. of Ed. Res.*, 8, 150-155.

Gray, P. L. and Marsden, R. E. (24-25). "The Constancy of the Intelligence Quotient—Further Results." *Brit. J. of Psych.*, 15, 169-173.

Gray, P. L. and Marsden, R. E. (26-27). "The Constancy of the Intelligence Quotient—Final Results." *Brit. J. of Psych.*, 17, 20-26.

Greene, K. G. (28). "The Influence of Specialized Training on Tests of General Intelligence." Chapter 21. Part I. *Twenty-seventh Yearbook. N. S. S. E.*, pp. 421-428.

Haggerty, M. E. (23). "Intelligence Examination Delta 2." *J. of Ed. Psych.*, 14, 257-277.

Henmon, V. A. C. and Burns, H. M. (23). "The Constancy of Intelligence Quotients with Borderline and Problem Cases." *J. of Ed. Psych.*, 14, 247-250.

Herring, J. P. (24). "Reliability of the Stanford and the Herring Revision of the Binet-Simon Tests." *J. of Ed. Psych.*, 15, 217-223.

Hildreth, G. (26). "Stanford-Binet Re-tests of 441 School Children." *Ped. Sem.*, 33, 365-386.

Hollingworth, L. S. (20). *The Psychology of Subnormal Children.* New York.

Hopkins, L. T. (24). *The Intelligence of Continuation School Children.* Cambridge University Press.

Hunsicker, L. M. (25). *A Study of the Relationship between Rate and Ability.* T. C. Contribs. to Educ., No. 185, Teachers College, Columbia University, N. Y.

Hunt, T. (28). "The Measurement of Social Intelligence." *J. of Appl. Psych.*, 12, 317-334.

Hurlock, E. B. (25). "The Effect of Incentives upon the Constancy of the I.Q." *Ped. Sem.*, 32, 422-434.

Johnson, B. J. (25). *Mental Growth of Children.* New York.

Jones, V. A. (27). "A Study of the Non-Verbal Nature and Validity of Myers' Mental Measure." *J. of Ed. Res.*, 16, 203-209.

Jordan, A. M. (23). "The Validation of Intelligence Tests." *J. of Ed. Psych.*, 14, 348-366 and 414-428.

Kelley, T. L. (28). *Crossroads in the Mind of Man.* Stanford University Press.

Keys, N. (28). *The Improvement of Measurement through Cumulative Testing.* T. C. Contribs. to Educ., No. 321. Teachers College, Columbia University, N. Y.

Kohs, S. C. (23). *Intelligence Measurement.* A Psychological and Statistical Study Based upon the Block-design Test. New York.

Lincoln, E. A. (27). "The Reliability of the Stanford-Binet Scale and the Constancy of Intelligence Quotients." *J. of Ed. Psych.*, 18, 621-626.

Lufkin, H. M. (21). "Report of the Use of the Army Alpha Test in Rural Schools." *Sch. and Soc.*, 13, 27-30.

Madsen, I. N. (24). "Some Results with the Stanford Revision of the Binet-Simon Tests." *Sch. and Soc.*, 19, 559-562.

Miller, W. S. (24). "The Variation and Significance of Intelligence Quotients Obtained from Group Tests." *J. of Ed. Psych.,* 15, 359-366.

Morgenthau, D. R. (22). "Some Well-Known Mental Tests Evaluated and Compared." *Arch. of Psych.,* No. 52.

Nettles, C. H. (26). "A Study on the Constancy of the I.Q." *Educ. Res. Bull.,* Los Angeles Schools, Vol. VI, No. 2, pp. 9-10.

Piéron, H. (26). "The Problem of Intelligence." *Ped. Sem.,* 33, 50-60.

Pintner, R. (24). "Results Obtained with the Non-Language Group Test." *J. of Ed. Psych.,* 15, 473-483.

Pintner, R. (25). "Group Tests after Several Years." *J. of Ed. Psych.,* 16, 391-395.

Pintner, R. (26). "An Empirical View of Intelligence." *J. of Ed. Psych.,* 17, 608-616.

Pintner, R. (27). "A New Intelligence Test for Survey Purposes." *T. C. Record,* 29, 18-23.

Pintner, R. and Upshall, C. C. (28). "Some Results of Social Intelligence Tests." *Sch. and Soc.,* 26, 369-70.

Proctor, W. M. (21). "Psychological Tests and Guidance of High School Pupils." *J. of Ed. Res.,* Monograph No. 1.

Randall, F. B. (27). "A Study on the Constancy of the I.Q." *Sch. and Soc.,* 26, 311-312.

Rogers, A. L. et al. (28). "The Constancy of the I.Q. and the Training of Examiners." *J. of Ed. Psych.,* 19, 257-62.

Rogers, M. C. (22). "Adenoids and Diseased Tonsils. Their Effect on General Intelligence." *Arch. of Psych.,* No. 50. New York.

Rugg, H. and Colloton, C. (21). "Constancy of the Stanford-Binet I.Q. as Shown by Re-tests." *J. of Ed. Psych.,* Vol. XII, pp. 315-322.

Rugg, L. S. (25). "Re-tests and the Constancy of the I.Q." *J. of Ed. Psych.,* Vol. XVI, pp. 341-343.

Sangren, P. V. (29). "Comparative Validity of Primary Intelligence Tests." *J. of Appl. Psych.,* Vol. XIII, pp. 394-412.

Shewrman, W. D. (26). "A Study of the Intelligence and Achievement of the June, 1925, Graduating Class of the Grover Cleveland High School, St. Louis." *Sch. Rev.,* Vol. XXXIV, pp. 137-146 and 219-226.

Spearman, C. (23). *The Nature of Intelligence and the Principles of Cognition.* London.

Spearman, C. (27). *The Abilities of Man.* New York.

Stenquist, J. L. (23). *Measurements of Mechanical Ability.* T. C. Contribs. to Educ., No. 130. T. C., Columbia Univ., New York.

Stern, W. (14). *The Psychological Methods of Testing Intelligence.* Trs. by G. M. Whipple. Baltimore.

Strang, R. (30). "Relation of Social Intelligence to Certain Other Factors." *Sch. and Soc.,* 32, 268-272.

Symposium (21). "Intelligence and Its Measurement." *J. of Ed. Psych.,* Vol. XII, Nos. 3 and 4, pp. 123-147, and pp. 195-216.

Teagarden, F. M. (24). *A Study of the Upper Limits of the Development of Intelligence.* T. C. Contribs. to Educ., No. 156. T. C., Columbia University.
Terman, L. M. (17). *The Stanford Revision and Extension of the Binet-Simon Scale for Measuring Intelligence.* Baltimore.
Terman, L. M. (19). *The Intelligence of School Children.* Boston.
Thorndike, E. L. (26). "On the Improvement in Intelligence Scores from Thirteen to Nineteen." *J. of Ed. Psych.,* Vol. XVII, pp. 73-76.
Thorndike, E. L. (23). "On the Improvement in Intelligence Scores from Fourteen to Eighteen." *J. of Ed. Psych.,* Vol. XIV, pp. 513-516.
Thorndike, E. L. (22). "Practice Effects on Intelligence Tests." *J. of Exp. Psych.,* Vol. V, pp. 101-107.
Thorndike, E. L. et al. (27). *The Measurement of Intelligence.* Bureau of Publications, Teachers College, Columbia University.
Thorndike, E. L., et al. (28). *Adult Learning.* New York.
Thurstone, L. L. (24). *The Nature of Intelligence.* New York.
Tilton, J. W. (26). *The Relation between Association and the Higher Mental Processes.* T. C. Contribs. to Educ., No. 218, Bureau of Publications. T. C., Columbia University.
Walters, F. C. (27). *A Statistical Study of Certain Aspects of the Time Factor in Intelligence.* T. C. Contribs. to Educ., No. 248.
Wells, F. L. (17). *Mental Adjustments.* New York.
Wentworth, M. M. (26). *Individual Differences in the Intelligence of School Children.* Cambridge, Harvard Univ. Press.
Westenberger, E. J. (27). "A Study of the Influence of Physical Defects upon Intelligence and Achievement." *Cath. Univ. Amer. Educ. Res. Bull.,* Vol. II, No. 9.
Whipple, G. M. (24). "Endowment, Maturity and Training as Factors in Intelligence Scores." *The Sci. Monthly,* Vol. XVIII, pp. 496-507.
Woodworth, R. S. (21). *Psychology.* New York.
Woolley, H. T. (26). *An Experimental Study of Children.* New York.
Worthington, M. R. (26). "A Study of Some Commonly Used Performance Tests." *J. of Appl. Psych.,* Vol. X, pp. 216-227.
Yerkes, R. M. [ed] (21). *Memoirs,* National Academy of Sciences, Vol. XV.

CHAPTER V

CRITERIA AND CONSTRUCTION OF INTELLIGENCE TESTS

The Type of Reaction.—Intelligence testing always consists of the testing of an individual's reactions. The efficiency with which an individual reacts is the measure of his intelligence. Some psychologists would restrict the type of reaction involved to relatively new situations or to situations of the problem-solving or reasoning or judgment type; others would not make any such restriction, claiming that habitual reactions or learned reactions may also be used as measures of intelligence. What an individual has already learned, what he has gained from experience, is a measure of his intelligence. If we stress the biological type of definition of intelligence, namely, the ability to react to relatively new situations, we shall seek items for our tests which are relatively new to all those whom we wish to measure. If we stress the behavioristic or empirical type of definition, namely, that intelligence may show itself in almost any kind of reaction, then we shall not be so particular as to the novelty of the items of our intelligence tests. From either point of view the most important thing is to make sure that our items are about equally novel to all those whom we wish to compare as to their intelligence, or that the individuals to be tested have had about the same opportunity to learn those items which we include in our tests. In other words novelty or non-novelty is always relative to the group we are testing, and whether we use the one kind of item or the other, we must attempt to make it suitable to the general background of the group. We cannot, therefore, make novelty or familiarity a criterion for the selection of items. The content of a test must be determined from some other point of view.

Validity.—When we ask how good a particular test of intelligence is, we are raising the question of its validity. How valid is this test as a test of intelligence? Is it really a test of intelligence or a test of something else? If it measures intelligence, how well does it do this? All of these questions refer to the validity of a test, and the validity of a test must be determined by means of some outside criterion of intelligence. Many criteria have been employed and we shall now discuss some of these.

(1) *Chronological Age.* This is one of the earliest criteria of intelligence and was employed by Binet. A test of intelligence should be passed by increasing percentages of children as we go from lower to higher ages. This criterion is based upon the hypotheses that intelligence increases as the child grows older, and that the distribution of intelligence among a random group is similar from year to year. Hence the score on a test or the percentage passing a given test should increase from year to year.

In his construction of the Stanford Revision of the Binet-Simon Scale, Terman (17) gives the percentages of children at each age passing the various tests. For example, Test VII, 3, Repeating Five Digits, is passed by the following percentages of children at the ages indicated:

Age	5	6	7	8	9
Per cent	34	59	74	83	93

And again, Test X, 6, Naming Sixty Words, shows the following percentages:

Age	8	9	10	11	12
Per cent	21	50	63	76	85

In a similar way, when dealing with group tests, we should expect the score to increase from age to age. Thus, Kuhlmann (28) uses this increase as a measure of the discriminative capacity of a test. When the percentage of increase in score from

one year to another becomes small, the test loses in discriminative capacity.

Kuhlmann (28) also uses the overlapping of scores between adjacent grades or ages as a measure of validity. This is of course another way of applying the criterion of age increase. If a test discriminates between one age and the next there will be little overlapping; if it does not there will be much or complete overlapping. Claparède (25) proposes the use of this measure of the amount of overlapping in order to distinguish between tests of age (or general intelligence) and tests of aptitude (independent of age).

This criterion of chronological age can never be used as the only criterion. It is a partial criterion only. It is of use after we have decided that a test is probably a test of intelligence, because the criterion alone cannot distinguish between intelligence traits and other non-intellectual abilities that mature as the individual grows. Height, weight, dexterity of movement and many other things increase as the child grows, but such things do not make up intelligence, nor are they good indicators of intelligence. This criterion, therefore, is of limited value and is not commonly used by psychologists to-day, although it is useful in helping to determine the relative discriminating values of a series of tests.

(2) *Known Groups.* This is also a criterion used by Binet in his early work. A test which discriminates between two groups of known intelligence is presumably a test of intelligence. If we give a test to a group of feebleminded and also to a group of normal children of the same age, then the normal children ought to do much better than the feebleminded. If not, if there is little difference between the two groups, then the test cannot be a good measure of intelligence. Similarly, we may use feebleminded and superior groups, or normal and superior groups. This procedure is really a very crude and coarse type of correlation. It is also allied to criteria to be mentioned presently, because the initial division of children into feebleminded and normal groups must have been made on

the basis of teachers' or physicians' judgments, or of school achievement, or of other psychological tests. In general, therefore, we may say that this criterion accepts the judgment of the world at large as to the intelligence of the groups, and uses this judgment as a criterion of intelligence. As a first rough measure of the goodness or badness of a test for intelligence testing purposes, it has proved useful. It has also been used by Hamid (25) as a method of analysis of tests to find out what qualities are desirable for intelligence tests.

(3) *Teachers' Judgments.* If a test is a test of intelligence it ought to correlate somewhat with the judgment of those best able to arrive at a judgment of the intelligence of a group of individuals. The judgment of teachers as to the intelligence of their pupils is, therefore, frequently used as a criterion of the validity of a test. Notice that we say the test should correlate "somewhat"; we do not expect anything like a perfect correlation. All individuals' judgments are fallible and we must not expect very much from the ordinary teacher. So many varied factors play a part in influencing the teacher, that we cannot expect more than a rough ranking of the pupils as to intelligence. Actual correlations between test scores and teachers' judgments of intelligence are as follows:

Author	*r*	*n*	*Test*
Ackerson and Estabrooks (27)	.50 to .60	89	Otis Group
Wilson (24)	.34	215	Terman Group
	.32	215	National
	.30	215	Otis
	.26	215	Northumberland
Viele (24)	.35 to .42	many	Pintner-Cunningham
	.33 to .38	classes	Detroit-Primary
	.42 to .51	in Grades	Dearborn Group
	.33 to .41	I and II	Cole-Vincent
Foran (24)	.64	65	Haggerty Delta I
	.61	65	Pressey Primary

Author	*r*	*n*	*Test*
	.76	65	Pintner-Cunningham
	.70	65	Detroit-Primary
Mitchell (24)	.42	1168	Stanford-Binet
Jordan (23)	.73	64	Otis Group
	.61	64	Army Alpha
	.68	64	Miller Group
	.66	64	Terman Group

These results may serve as samples of the magnitude of the correlations usually obtained. In general we expect the correlations to lie between .3 and .6. The variations depend upon the familiarity of the teacher with the pupils, the ability of the teacher to form such judgments, the range of talent among the pupils, and many other factors. Furthermore, we must remember that the correlation tends to rise if the judgment of the child's intelligence is the average rating of more than one teacher. Jordan's data were based on the average rating of each pupil by four teachers. If we are constructing a new test and find that it correlates between .3 and .6 on the average with teachers' judgments, we should be satisfied so far as this criterion of validity is concerned. We should be suspicious of its validity if the correlation dropped much below .3, particularly if other intelligence tests with the same pupils were giving correlations above .3.

(4) *School Achievement.* School achievement may be expressed by school marks allotted by teachers, or by scores on standard educational tests or by the rate of progress through the grades. Whenever we take any of these ratings as a measure of validity for our intelligence tests, we assume that the intelligent child will work more or less up to his capacity in his school work. In rate of progress we assume further that he will be allowed to move through the grades somewhat in accordance with his intelligence. These assumptions are on the whole true, so that we should expect positive correlations. There are, however, enough individual exceptions to prevent anything like

a perfect correlation between intelligence and school achievement.

Here are samples of correlations between intelligence tests and school marks copied from Symonds (27):

r	*Author*
.37	Ross
.51	Flemming
.52 (Binet)	Proctor
.38 (Group)	Proctor
.48	Jordan

Jordan (23) has collected the correlations for three tests found by numerous workers and his averages and ranges are given below:

	r Median or Average	*Range of r's*	*n*	*Grade*
Army Alpha	.38	.19-.52	26	High School
Army Alpha	.41	.22-.67	35	College
Otis Group	.66	.33-.91	40	Grades IV-VIII
Otis Group	.49	.31-.82	16	High School
Terman Group	.47	.30-.67	9	High School

Wilson (24) found the following correlations for four group tests given to the same children, ages 10, 11 and 12:—

Terman Group	.49
Northumberland	.42
National	.39
Otis	.39

These correlations with school marks are very much like the correlations with teachers' estimates of intelligence. Teachers are naturally influenced in the estimates of intelligence by the school achievement of their pupils. More objective than school marks are standard educational tests. Samples of correlations between general intelligence and educational tests are on facing page.

It will be noted that these correlations are in general some-

Author	*r*	*n*	*Grade*	*Intelligence Test*	*Educational Test*
Kelley (27)	.71	25	IV	Illinois	Stanford
	.66	60	VI-VIII	National	Stanford
	.79	156	VI	Otis S. A.	Stanford
Gates and La Salle (23).....	.88	75	III-VI	National	Thorn-McCall Reading
	.91	75	III-VI	National	Woody Arithmetic
	.85	75	III-VI	National	Ayres Spelling
Witty (29)	.63	522	IV-VI	Binet	Stanford
	.78	522	IV-VI	Multimental	Stanford
Gates (22).	.22	20	IV	Myers	Composite
	.61	20	IV	Dearborn	Composite
	.65	20	IV	National	Composite
	.46	20	IV	Haggerty	Composite
	.45	20	IV	Illinois	Composite
	.69	20	VIII	Otis Advanced	Composite
	.58	20	IX	Terman	Composite

what higher than those with teachers' estimates and school marks. This may be due to the better rating of achievement by means of standard educational tests. It is also due to the community of function between intelligence and achievement tests. They sometimes measure very similar reactions, as, for example, when an intelligence test contains an arithmetic reasoning sub-test and a composite educational test contains a similar sub-test. Kelley (27) estimates this community of function between very verbal intelligence tests and achievement batteries like the Stanford to be about 90 per cent. Non-verbal intelligence tests have in all probability a very much smaller percentage of community of function. They do not correlate very highly with educational achievement, as we see from the example of the Myers Test above.

The use of educational tests as measures of validity for intelligence tests is, therefore, of limited value. An intelligence test should correlate fairly well with educational achievement, but we cannot use an educational achievement test as the sole criterion. If we do so, we could build our intelligence test in such a way as to reach an almost perfect correlation simply by using more and more of the content of the educational test. This, of course, might be a perfectly valid test for a restricted group of school children all taught exactly the same things, but it would become less and less valid as an intelligence test for children in schools differing widely in their curricula. For comparisons of intelligence between such children we must get away from the specific content taught in the schools as much as possible. This is, of course, the only difference in content between intelligence and educational achievement tests, namely, that the former contain a wider and more general sampling of stimuli than do the latter, in order that we may compare children of different school experiences and thus rise to a common background beyond the specific things taught in each particular school. So we select opposites, analogies and number completions for our intelligence test and we leave spelling and writing and geography for the educational achievement test.

(5) *Other Tests.* Another measure of validity is the correlation with a known test of intelligence. If the Stanford-Binet is accepted as a valid measure of intelligence, then a new test offered as a measure of intelligence should correlate positively with the Stanford-Binet. All correlations of one intelligence test with another may from this point of view be used as validity coefficients. The only trouble is to know which of the two tests to take as the criterion and which the one whose validity is being measured.

If we are perfectly convinced that there is but one kind or type of general intelligence and that the Binet test is the best measure of it, then we may naturally take the Binet as a valid criterion. If we believe that there are different kinds of intel-

ligence, positively but not perfectly correlated, then we shall expect a fair correlation of other intelligence tests with the Binet as a criterion. Here are a few sample correlations of other tests where the Stanford-Binet has been used as a kind of validity criterion:

Author	*r*	*Test*	*Comments*
Herring (24)	.98	Herring-Binet	Wide Age Range
Avery (24)	.77	Herring-Binet	Grade I
McAnulty (28) ..	.63	National	3 Grades
	.73	Terman Group	3 Grades
Heckman (24) ...	.77	Detroit Primary	Grade I
	.49	Pintner-Cunningham	Grade I
Dougherty (28) ..	.80	Pintner-Cunningham	Kindergarten
Pintner (27)	.73	Pintner-Cunningham	229 Cases
	.88	Pintner-Cunningham	72 Cases
Lincoln (28)	.85	Dearborn	55 Cases

We need not take the Binet Test as the criterion. We may take some other intelligence test, but unless this other test has been adequately validated, our correlations only show the similarity between the two tests and the validity of either is still an open question.

(6) *Combinations of These.* Any one of the five methods of determining validity has its drawbacks, therefore some workers have used a combination of these hoping thereby to arrive at a better criterion. Thus McCall (26) uses a criterion made up of teachers' judgments, Binet, group intelligence tests and measures of educational achievement. With this composite criterion his Multi-Mental Test correlates .93 for 90 cases in Grades III to VIII. Liu (22) has used the most elaborate and probably the best criterion for estimating intelligence. It was made up of (1) age, (2) school marks, (3) school progress, (4) teachers' estimates of intelligence and (5) composite test scores of five group intelligence scales. The elements entering into the criterion were carefully weighted. Some of the correlations with this criterion are:

Dearborn Group Test	.80	235 cases
Pintner Non-Language Test	.78	235 cases
Army Beta	.75	235 cases
Myers Mental Measure	.65	235 cases
Pressey Primer Scale	.58	237 cases

Very frequently a composite of several intelligence tests is used, as in the study of McGraw and Mangold (29), where each of ten tests is correlated against the composite of ten. They find by this method that of ten primary tests Otis Primary correlates .79 with the composite; Dearborn Primary .78; Pintner-Cunningham .77; and so on. In this same study there is a suggestion of still another criterion of validity, namely the correlation between a test rating and another rating of intelligence several years later. The assumption here is that intelligence is fairly constant and therefore that test is the most valid which will give a rating now that will correlate highest with a rating several years later. This, of course, assumes a valid rating several years later. McGraw and Mangold gave the Otis Group Test five years later and found the following correlations with the intelligence ratings given five years previously to the children when in Grade I:

Dearborn Primary	.60
Pintner-Cunningham	.55
Otis Primary	.53
Haggerty Delta 1	.53

From this survey of the various methods of determining the validity of an intelligence test, we can see that there is no one method that is infallible. Each single method is open to objections. The psychologist must use as many of these methods as he can. The construction of a composite criterion is undoubtedly the safest plan.

Reliability.—If we give a test that we know is a valid measure of intelligence, we want also to be sure that the test is a reliable measure of intelligence. This means that if we repeat the measure we would get the same result. An un-

reliable test would not give us the same result. A reliable yardstick is one that will give us the same measure for the length of this table if we measure the table twice. An unreliable yardstick would be one that would shrink two or three inches between our measurements. The coefficient of correlation between two forms of a test is the usual measure of reliability. If there is only one form, the correlation between two random halves of the test, taking even and odd items, is generally used, after being corrected by the Spearman-Brown formula. A somewhat less adequate measure of reliability is the correlation between two trials of the same test. This correlation is generally higher than the real measure of reliability.

Since all coefficients of correlation are dependent upon the variability of the group, no test has one reliability coefficient. It will have one for a narrow population as in a single grade, it will have a higher one for a population of several grades. It is therefore difficult to compare the reliabilities of one test with another unless these are computed on the same or similar populations. Some reported reliabilities are as follows:

Test	*r*	*n*	*Range*	*Author*
Stanford-Binet . .	.93	428	C. A. 3–15 years	Terman
Pintner-Cunningham.	.93	147	Grades I and II	Cureton
Kuhlmann-Binet.	.82	393	C. A. 18–54 months	Goodenough
National A.	.92	232	Grades IV-VIII	Symonds
National B.	.95	232	Grades IV-VIII	Symonds
National A.	.70		Estimate for single grade	Kelley (27)
National B.	.75		Estimate for single grade	Kelley (27)
Haggerty 2	.6		Estimate for single grade	Kelley (27)
Multi-Mental . . .	.94		Grades III to IX	McCall
Terman Group . .	.89		Grade IX	Terman
Pintner-Non-Language.	.79	201	Grades IV-VI	Pintner

The best source for the reliabilities of tests is Kelley (27). He has tried to assemble all the available data on reliabilities of published tests of intelligence and achievement. He points out that the requirements as to reliability vary markedly with the purpose for which a test is used. Tests need not have reliabilities much above .5, if we only wish to use them for the measurement of groups. In such cases we do not use the scores of individuals, but only the average scores of the groups. When, however, we wish to use individual scores to estimate the probable future standing of the individual, the reliability of the test should be about .9. Few of our present group tests of intelligence attain this degree of reliability for a single grade range, and therefore a single group test must be used with caution for individual diagnostic purposes.

The Ways of Expressing Intelligence Ratings.—There are many different ways in which we may express the intelligence rating of an individual, and we shall discuss some of those most commonly used at the present time.

1. *Mental Age.*—One of the most common and serviceable means of designating the intelligence of a child or of adults of inferior intelligence is by what is known as the Mental Age. This method originated with the Binet Scale, but is now used with any test standardized on children of various ages. If a child is said to have a certain mental age, we mean thereby that his performance on the test is equal to the average performance of a fair sampling of children of the same chronological age. Thus, if average five-year-olds are able to pass certain tests, a child, regardless of his chronological age, who is able to pass these same tests and unable to pass any higher ones is said to have a mental age of five years.

The accuracy of any mental age depends, of course, upon the thoroughness of the standardization of the scale from which the mental age is computed. It is, therefore, always necessary to keep in mind the scale upon which the child was tested and it is dangerous to compare the mental ages of children tested by different scales. The original Binet Scale was found to be

too easy at the lower end and too hard at the upper end for average American children. A child having a mental age of three or four on this scale cannot, therefore, be properly compared with a child whose mental age is three or four as tested by the Stanford Revision of the Binet Scale. It is, therefore, always necessary to keep in mind the scale from which the mental age has been computed. At the present time the Stanford Revision would seem to be the best standardized scale and it is the one most widely used. It may not be absolutely accurate and there is a possibility that the higher ages (12, 14, 16) are too hard for the average child of those ages; nevertheless constant use of the scale gives us a familiarity with its meaning and something like a conventional significance is being attached to the different mental ages on the Stanford Revision. They are beginning to stand for specific degrees of intelligence, even although they may not in every case actually measure the average ability of the age in question.

Although the concept of mental age is extremely useful and readily understandable in practical work, it has one grave drawback. This becomes at once apparent when dealing with adults who are not backward, or with very superior children. Because mental age means the average or median performance of a fair sampling of the equivalent chronological age group, we cannot really have a mental age above 14 or 16, if mental growth stops at these ages. In other words, a mental age of 12 is equivalent to the average performance of twelve-year-olds; a mental age of 13 to the average performance of thirteen-year-olds; a mental age of 14 to the performance of fourteen-year-olds; a mental age of 15 to the performance of fifteen-year-olds; a mental age of 18 to the performance of eighteen-year-olds; a mental age of 20 to the performance of twenty-year-olds; and so on to any age. Now if mental growth ceases at 14, then the average performance of 14-, 15-, 16-, 18-, 20-, 30-year-olds must be identical. Therefore, we cannot go above a mental age of 14. Any individual exceeding the score of the average fourteen-year-old has a mental age above 14, but that

is really all we can say about his mental age. In order to assign mental ages to such individuals so as to be able to calculate I.Qs., many workers have assumed norms for higher ages above 14 or 16. In doing this, however, we should remember that these so-called mental ages are not based upon the strict meaning of mental age, as the average performance of the like chronological age group. Such mental ages are pure fictions and their comparability with true mental ages is doubtful. The mental age method is, therefore, limited in scope, and it is for this reason that other methods of rating intelligence, such as the percentile method or the standard deviation method, are coming into use, particularly in dealing with older individuals.

Mental age is an absolute measure of the child's intelligence. The relative intelligence of the child can only be estimated by comparing his mental with his chronological age. The easiest comparison is to note how many years above or below his chronological age he may be. This, however, is very unsatisfactory because a difference of a certain number of years has a totally different significance at different periods of life. A four-year-old who is two years retarded mentally is much more seriously deficient than a twelve-year-old who is also two years retarded mentally. This is because the growth of intelligence is much more rapid at the earlier ages and gradually becomes slower and slower. Our previous discussion on the growth of intelligence has made this clear. Owing to this difficulty we must, therefore, adopt some other method of rating a child's intelligence, if we wish to compare the intelligence of children of different chronological ages.

2. *The Intelligence Quotient.*—The intelligence quotient gives us the relative degree of intelligence disregarding chronological age. It is the ratio of the mental and chronological ages.

The formula is $I.Q. = \frac{M.A.}{C.A.} \times 100$. The multiplication is simply to get rid of decimals. Children testing exactly at age

will all have I.Qs. of 100, those below age will have I.Qs. less than 100, and those above will have I.Qs. greater than 100. The Intelligence Quotient thus gives us an intelligence rating that makes possible direct comparison of children of different chronological ages. It is the most serviceable value when dealing with age-scales.

A study of the distribution of the I.Qs. of a large number of children, such as has been made by Terman (17), shows a fairly constant distribution at each age. In his combined distribution for all ages from five to fourteen he gets the following results:

The	lowest	1	per	cent	go	to	an	I.Q.	of	70	or	below.
"	"	2	"	"	"	"	"	"	"	73	"	"
"	"	3	"	"	"	"	"	"	"	76	"	"
"	"	5	"	"	"	"	"	"	"	78	"	"
"	"	10	"	"	"	"	"	"	"	85	"	"

and so on.

The	highest	1	per	cent	reach	an	I.Q.	of	130	or	above.
"	"	2	"	"	"	"	"	"	128	"	"
"	"	3	"	"	"	"	"	"	125	"	"
"	"	5	"	"	"	"	"	"	122	"	"
"	"	10	"	"	"	"	"	"	116	"	"

and so on.

The diagnostic significance of the different Intelligence Quotients is generally expressed somewhat as follows:

Intelligence Quotients	*Classification*
0-69	Feebleminded
70-79	Borderline
80-89	Backward
90-109	Normal
110-119	Bright
120-129	Very Bright
130 and above	Very Superior

Several methods for the rapid calculation of Intelligence Quotients are available. Tables for all I.Qs. within the mental age range from three to seventeen years and the chronological age range from five to seventeen years have been prepared by Inglis (22). Toops and Pintner (18) have devised a chart for reading off Intelligence Quotients. Yepsen (22) has devised a slide rule of the conventional type called the Vineland Slide Rule, while Kohs has constructed a circular slide rule, called the Reed I.Q. Slide Rule. The tables are undoubtedly the most accurate, but the other devices are rapid and may be useful where extreme accuracy is not necessary.

3. *Point Scales and Their Coefficients.*—If, instead of an age scale, we are using a scale on which a child is said to score so many points (see description of Yerkes-Bridges Scale in Chapter VI), we obtain as a measure of his intelligence a score of so many points. We may then refer to the norms and turn this into a mental age and from this derive an I.Q. as explained above. Instead of this procedure, however, we have the possibility of comparing the number of points obtained with the number of points normal for the age in question. The ratio of these two values gives a coefficient which has been called the Coefficient of Intellectual Ability (C.I.A.) or the Coefficient of Mental Ability (C.M.A.). This coefficient expresses the percentage of points gained with reference to the norm for the age in question. It cannot be directly compared with the I.Q. and the diagnostic significance must be empirically determined for the scale with which it is used. So far it has only been used with the Yerkes-Bridges Scale, and we shall give the diagnostic significance of the coefficient, as far as has been determined, when the scale itself is discussed.

4. *The Median Mental Age.*—When dealing with a number of standardized tests, each of which is significant for a wide range of ages, a very useful method of intelligence rating is the median mental age. Suppose we have fifteen tests showing significant scores or time values at most ages from five to fourteen, we can turn any particular child's score into the equiva-

lent mental age for each test. We then have fifteen mental ages descriptive of his performance on the tests. The median of these fifteen mental ages gives us a mental age which can be regarded as the most likely mental age of the child. This mental age can then be converted into an I.Q.

This method of evaluating intelligence has been used by Pintner and Paterson with their Performance Scale. (See Chapter VI.) It has also been more recently applied by Kuhlmann (27) to a battery of group tests standardized by him. (See Chapter VII.) The value of the method lies in the flexibility allowed the examiner with reference to the number of tests used in any particular case. Indeed a mental age of this kind can be obtained from any conglomeration of properly standardized tests which cover a sufficiently wide range of ages. Furthermore, the list of mental ages for the individual tests gives the examiner a picture or profile of the child's performance showing at a glance the tests or types of test in which the child excells or fails.

5. *Percentile Ratings.*—A type of rating that is becoming more and more popular among mental testers is the percentile rating. In this case the individual's performance is compared with the performances of all the individuals in a particular homogeneous group. A six-year-old is compared with a group of six-year-olds; a ten-year-old with a group of ten-year-olds; a college student with a group of college students or, if we like, with a group of miscellaneous adults. The reference always is to a group of which the individual is in some way a member. We do not compare a five-year-old's performance with the average performance of children of different ages and find that it is like the average performance of six-year-olds or seven-year-olds. We compare his performance with five-year-olds only and find out whether it is equal to the performance of 10 per cent, 20 per cent, 50 per cent or any per cent of five-year-olds. This percentile value is the measure of his ability and it is easy to interpret. It ranks him immediately on a scale of one hundred. It tells us how many are above or below him, and we can then

describe his ability in any way we wish in accordance with whatever sub-division of groups we have carried out on our normal curve of distribution.

The percentile method can be used with a group of tests each of which has been standardized according to percentiles. In this case we shall have a percentile rating for each test and the median or average of these percentiles will be the intelligence rating of the child.

This method has been used by Woolley (15) in her tests for adolescents and also by Pintner (18). The latter shows the necessity for "super-percentile" tables for the accurate interpretation of an average or median derived from several percentile ratings. On the other hand we may have a series of tests each of which is assigned a score, and the total score will then be converted into a percentile which will give the individual's rating.

The percentile method has found particular favor with group tests. This is probably owing to the fact that it demands a large number of cases in order to give a reliable standardization. It is unquestionably one of the best methods of expressing an intelligence rating and its use will become more common in the future. With percentile norms for each age group we avoid the necessity for comparing subjects of one chronological age with those of another, a procedure which Chapman and Dale (22) have criticized. They question the practice of rating a precocious nine-year-old on tests devised for twelve-year-olds. "A child of 8 and a child of 12 cannot be compared." They believe that finally the mental age and the Intelligence Quotient must be abandoned.

6. *Standard Deviation Measures.*—Probably the best intelligence rating and the one that will be extensively used in the future is a rating which expresses the standing of the subject with reference to the standard deviation of the group with which he is to be compared. The advantage of any standard deviation unit over a percentile rating is that the standard deviation units are equal at all points of the difficulty scale.

The differences between any two percentile ratings are not equal, because differences increase as we go out from the median either upwards or downwards.

The individual's standing on a test is expressed in terms of the standard deviation of the group. This rating is called the standard score and it is

$$Z = \frac{X - M}{\sigma},$$

where X is the individual's score, M the mean score of the group and σ the standard deviation (Kelley, 23). McCall's (22) T Scale is a variation of this method. The T unit is one-tenth of the standard deviation of an unselected sampling of twelve-year-old children. As we shall see below, McCall has used this method to good effect in the combination of intelligence and achievement ratings.

7. *Combined Educational and Intelligence Ratings.*—Although it is not the purpose of this book to discuss educational tests, we must call attention to the devices which are coming into use for evaluating educational attainment in terms of intelligence. In general two types of devices are used. The first is based upon standard deviation ratings, the second upon quotient ratings.

Pintner and Marshall (21) employed the mental index (a standard deviation rating) for the Pintner Non-Language and Educational Survey Tests. The "Difference" between these two indices gives a measure of achievement in terms of intelligence. McCall (22, 23) has extended his T Scale method to cover both intelligence and educational tests. His *F*, or efficiency rating, is the difference between a T score (Te) on an educational test and a T score (Ti) on an intelligence test. Many workers have used the straight standard score technique for both educational and intelligence tests. The difference between such standard scores gives a measure of educational attainment in terms of intelligence.

The other type of combination of intelligence and educa-

tional ratings was suggested by Franzen (22). He calculated an educational age (E.A.) in the same manner as mental ages (M.A.) had been calculated. Thus we may have an arithmetic age, spelling age, and the like as a measure of a child's ability and these ages are determined from educational tests in the same way as mental ages are determined from intelligence tests. The E.A. (educational age) divided by the C.A. (chronological age) gives an E.Q. (educational quotient), just as an M.A. divided by a C.A. gives an I.Q. The ratio of E.Q. to I.Q. gives an accomplishment quotient or A.Q. The ratio of E.A. to M.A. gives of course the same thing. This Accomplishment Quotient is a measure of the educational attainment of the child in terms of his intelligence.

The A.Q. technique has been widely used in educational research. As a general rule we find a negative correlation between I.Q. and A.Q. Thus Popenoe (27) reports a correlation of —.45, and Beeson and Tope (24) —.46. In other words the children of high I.Q.'s do not attain as high E.A.'s as M.A.'s. One obvious reason for this is that such children are not able to do so in the ordinary school. They are generally to be found in grades much below their intellectual level. Furthermore, as Goodenough (25) points out, they have been in school for a much shorter period of time, for example, a child with I.Q. 140 and M.A. 14 has spent 7 semesters in school, whereas a child of I.Q. 100 and M.A. 14 has spent about 15 semesters in school.

There is much doubt as to the value of the A.Q. for school purposes. It is a derived score depending upon at least two standardizations, one of a mental test and the other of an educational test. Unless these two standardizations are comparable, i.e., based upon the same kind of sampling, we have no means of knowing how to interpret the A.Q., derived from the I.Q. and E.Q. of our two standardizations. The best results with the A.Q. technique would be obtained by the use of educational and intelligence tests standardized on the same populations. Popenoe (27) has attempted to measure the reliability

of the A.Q. by giving two repetitions of the intelligence and educational tests and calculating two A.Q.'s for each child. The correlation is .28 for 604 cases in grades III to VIII. If this should prove to be the usual correlation for most intelligence and achievement tests, then the A.Q. technique cannot be used to advantage at present in practical school work.

STANDARDIZATION

By standardization we mean the establishment of a certain definite method of giving a test and the establishment of adequate norms for the interpretation of the results. The former may be called standardization of procedure and the latter standardization of results.

Standardization of Procedure.—It is imperative in giving a test to give it in as nearly as possible the same manner in which the test was originally given, that is, of course, if we wish to compare our results with those of the originator or if we wish to measure the intelligence of a child by means of some recognized test. Many conflicting differences in the results of different workers may be traced directly to slight differences in procedure. Just how significant a slight change in procedure is going to be can never be foretold. We must actually try out the difference before we can tell what influence it will have. Therefore, when we are using a standard scale for practical results, we must be very careful to follow exactly the procedure as laid down by the author. We may think it would be better to present a test in this or that fashion, and it may really be so, but we must remember that if we do so, we may be changing the test and we cannot, therefore, use the norms which the author of the scale has arrived at. To be sure, by making a certain change, we may make a much better test, but if we do that, we must remember that it is a *new* test and before using our results we must prepare *new* norms by which to interpret them.

All of this sounds trite and commonplace to the psychologist,

but it is something that cannot be too strongly impressed upon the beginner in intelligence testing. The teacher in particular, who is learning to give intelligence tests, must keep this in mind, because her whole attitude in the class room is one of adapting the presentation of a subject to her pupils and in doing so the good teacher turns it around and presents it from many different angles, so as to reach many different grades of intelligence. She shows her initiative by doing this and she is quick to perceive when a child does not grasp the point, and ready to give the subject just a little turn to bring it within the focus of his intelligence. It often happens, therefore, that in presenting a mental test she grasps at once the difficulty that the child is experiencing and knows what slight modification of the presentation would bring it within his range, and the temptation to do this is great. The teacher must, however, resolutely withstand any such temptation, because she must remember that, if her slight modification had been made and so given to all the children upon which the test has been standardized, the norms would have been very different from what they now are. The teacher or examiner must remember that she is not teaching but testing the child.

Standardization of Results.—By this we mean the establishment of adequate standards or norms by means of which the result of testing any individual case can be properly interpreted. It is, of course, necessary to give our test or tests to a sufficiently large number of cases before we can be sure of having arrived at a true average or median. Many tests are of doubtful practical value, because they have not been tried with enough cases to warrant any confidence in the averages or medians secured. It is difficult to say just how many cases must be tested before a satisfactory standardization can be reached. Perhaps the most satisfactory method to pursue is to work out the age or percentile norms at different stages in the standardization, and note whether the averages, medians or percentiles fluctuate much or little as each new group of cases is added. After testing, say, fifty children, work out the

norms. Add another group of fifty or so, and work out the norms again; and so on with additional increments, until a general tendency for the norm to remain stable is noticed. The presumption then is that the addition of more cases will not materially alter the norm and we may regard our standardization as complete.

In this work of standardization our constant endeavor is to get a fair sampling of cases at each age. By a fair sampling is meant a group which will represent in proper proportions all the different types of intelligence contained within the large group. Obviously we must be content with a sampling of any group. To determine the intelligence of the normal American citizen, we cannot test all the one hundred million inhabitants of the United States. To arrive at the ability of six-year-olds on some given test, we cannot test all the six-year-olds in the United States. Nor is it really necessary to do this. All that we need is a fair sampling of six-year-olds. Theoretically it is very difficult to decide as to what a fair sampling of any group is, and we might enter into a lengthy critique of almost all standardized tests from the point of view of a fair sampling and it would not be difficult to find fault with most standardizations. Such theoretical exactness must not be allowed to stand in the way of the practical work of gathering results to arrive at standards. What the psychologist generally does is to use the ordinary school child, and if he gets a fair mixture of city and rural children, the probability is that he will obtain a fair sampling.

In using elementary school children for the lower ages the psychologist realizes that the most defective cases have been eliminated, because the idiots are almost always in institutions or at home; similarly in the more progressive states the majority of imbeciles are likewise eliminated from the school. On the other hand the brightest children are not automatically eliminated in this manner. Whether this discrepancy between the lower and upper end of our distribution will seriously injure the ordinary standardization is doubtful.

When we come to the higher ages, say thirteen and above, the difficulty of getting a fair sampling is greatly increased. The larger number of children leave school at the age of fourteen and a great many drop out even before that age. Those that remain belong, in the main, to the wealthier classes and these classes are endowed on the average with more intelligence than the poorer classes. Furthermore, children of mediocre or inferior intelligence, of whichever class, tend to drop out of school sooner than their better endowed fellows. Such children find they cannot keep up the pace set in the High School, they become uncomfortable and restless, and leave school in order to find a more congenial environment. All standardizations for these higher ages, which are based upon school children entirely, lead in all probability to norms that are too high for the population in general. In the Stanford Scale a splendid attempt was made to arrive at average adult intelligence and it was thought to correspond to the ability of sixteen-year-old students (mostly High School). The testing of thousands of average adults in the army leads us to believe that the Stanford tests for age fourteen are more nearly adequate measures for the average American adult. This illustrates the difficulty of getting a fair sampling of children at each age, particularly in the higher ages.

The further we progress in intelligence testing the more clearly do these difficulties of obtaining a random or unselected group of cases emerge. The wide application of intelligence tests to all types of people and in all types of communities has revealed individual differences far greater than were thought to exist. Different sections of a city differ greatly in intelligence and so do different rural communities. We find differences in the average intelligence of the inhabitants of different small towns and the same is probably true of cities, and of states and larger sections of the country. The existence of such differences need not dismay us nor make us give up as hopeless the standardization of intelligence tests, but the knowledge that they exist should at least keep us from being dogmatic and

from regarding any scale as being a complete and final measuring rod of intelligence.

Having obtained results from as fair a sampling as possible, they will now be used by the psychologist in different ways according to the type of scale he is standardizing.

Age Scales.—For age scales of the Binet type, where each test is either passed or failed, the results will consist of the percentage of children at each age who have passed a given test. The test must fulfill the criteria we have laid down above and is said to be standard for the age at which about 75 per cent (roughly sixty to eighty per cent) pass. The assumption is that the middle 50 per cent and also the upper 25 per cent will pass the test, making in all about 75 per cent. Actually in the Stanford Revision we find these percentages running from about 60 to about 80.

Point Scales.—Here our results will consist of the total number of points scored by each child and the norms will be expressed by the average or median number of points for each age. This average must show a steady rise from age to age.

Median Mental Age.—Here the results will consist of some kind of a score or time value. The median score for each age is determined and a test has discriminative value for that range of ages which shows a progressive increase in score.

Percentile Ratings.—The scores of all the cases at each age are arranged in order of merit and percentiles calculated. We may calculate as many different percentiles as we wish. The more percentiles we calculate, the larger the number of cases must we possess in order to have a reliable standardization. Because of the fact that we take account of extreme low and high values, which are later to be used as diagnostically significant, the percentile method requires more cases for an adequate standardization than do the other methods. Each percentile should show a progressive rise in score as we go from age to age.

The Construction of Intelligence Tests.—The construction of tests has become a complicated and technical matter.

There is no one way in which a test must be constructed. There will be differences according to whether the test is to be an individual or a group test. The construction of an individual test of the Binet or Performance type is a very long and tedious task. Group tests can be more quickly constructed. We shall give a general description of the most important steps in group test construction.

(1) *Decision as to General Characteristics.* A decision must be made as to the kind of test to be constructed, whether a long test for individual diagnosis or a briefer test for survey purposes. Furthermore, the range of ability must be determined upon; whether the test is to be for kindergarten and first grade, or for the upper elementary grades or for the high school, and so on.

(2) *Assembly of Items.* First the kinds of items will be chosen. Some tests have one kind of item, e.g., the McCall Multi-Mental Scale; others have several kinds, such as Opposites, Analogies, Number Completion and so on. More kinds of items can be chosen than the experimenter may wish finally to retain. A new type of item, that has not previously been used in intelligence tests, may be included.

Having determined the kinds of items, the items themselves are assembled. About twice as many as may finally be used are collected, and they are arranged in a guessed order of difficulty. Or the order of difficulty is made more carefully by expert judgment. The combined judgments of several experts will save a great deal of time later on.

(3) *First Try-Out.* These long preliminary tests are given to several hundred children, the number being dependent upon the range of ability the test is supposed to cover. The results from this try-out may be used to settle how many sub-tests will be used, and whether any new type of sub-test is valid enough to be included in the final test. At this stage items will be validated by some kind of criterion (Brigham, 28). The items will also be arranged in order of difficulty by percentages of those passing or by standard deviation values.

Items that are not valid or items of equal difficulty not required will now be discarded. If more than one form is planned, it is at this stage that items will be chosen for the different forms, items of equal difficulty as far as possible being put into the various forms, so as to make the total difficulty of each sub-test of each form equal.

(4) *Second Try-Out.* All the forms of the test will now be given to another sampling of pupils covering the whole range of ability for which the test is planned. Each pupil will be given all the forms, but the forms will be rotated, some pupils getting Form A and then Form B, and an equal number Form B and then Form A; and similarly with other forms, should there be more than two forms of the test. The results of this try-out are then used to test the equality of the forms and to make such changes in items as may be necessary to insure this equality. If the preliminary work has been well done, there should not be many changes. If this is not the case, however, all the items may have to be tabulated, difficulty values again determined and new forms arranged. An additional try-out would then be necessary to check up on these new forms.

During the second try-out it is generally possible to determine time limits for such tests as require them, and also to decide finally upon the instructions for giving and scoring the test.

(5) *Third Try-Out.* The test should now be printed in its final form and given to a great many pupils. The results from this try-out will be correlated against any criteria of validity that may be obtainable, so as to obtain validity coefficients for the final test. Many pupils will be given two or more forms of the test in order to calculate reliability coefficients. These correlations should refer to the whole range of ability which the test covers, as well as to the reliability of single grade groups. In addition the probable error of a test score should be calculated (Ruch and Stoddard, 27).

The results of this final try-out will also be used for the

establishment of norms. Such norms will probably be age norms, showing the increase in score by age for the calculation of M.A.s. The equivalent M.A. for each score is desirable here. Percentile norms for age groups may also be calculated. A table giving raw scores and T scores is also desirable, now that the T score is being frequently used (McCall, 23).

REFERENCES

Ackerson, L. and Estabrooks, G. H. (27-28). "On the Correlation of Intelligence Test Scores with Imputed Intelligence." *Brit. J. of Psych.*, Vol. 18, 455-459.

Avery, G. T. (24). "Comparison of Stanford and Herring Binet Revisions Given to First Grade Children." *J. of Ed. Psych.*, 15, 224-228.

Beeson, M. F. and Tope, R. E. (24). "The Educational and Accomplishment Quotients as an Aid in the Classification of Pupils." *J. of Ed. Res.*, 9, 281-292.

Brigham, C. C. (28). *Third Annual Report.* Commission on Scholastic Aptitude Tests. College Entrance Board. New York.

Chapman, J. C. and Dale, A. B. (22). "A Further Criterion for the Selection of Mental Test Elements." *J. of Ed. Psych.*, 13, 267-276.

Claparède, E. (25). *Comment diagnostiquer les aptitudes chez les écoliers.* Flammarion, Paris.

Dougherty, M. L. (28). *A Comparative Study of Nine Group Tests of Intelligence for Primary Grades.* Baltimore, Johns Hopkins Press.

Foran, T. G. (24). "A Study of Some Group Intelligence Tests in the First Grade." *Cath. Educ. Rev.*, 22, No. 2.

Franzen, R. (22). *The Accomplishment Ratio.* T. C. Contribs. to Educ., No. 125, T. C., Columbia University.

Gates, A. I. (22). "The Correlations of Achievement in School Subjects with Intelligence Tests and Other Variables." *J. of Ed. Psych.*, 13, 129-139 and 223-235.

Gates, A. I. and La Salle, J. (23). "The Relative Predictive Values of Certain Intelligence and Educational Tests." *J. of Ed. Psych.*, 14, 517-539.

Goodenough, F. L. (25). "Efficiency in Learning and the Accomplishment Ratio." *J. of Ed. Res.*, 12, 297-300.

Hamid, S. A. (25-26). "Some Factors of Effectiveness in Mental (Intelligence) Tests." *Brit. J. of Psych.*, 16, 100-115.

Heckman, S. B. (24). "A Comparative Study of Group Intelligence Tests Applicable to Children of Kindergarten Age." New York Society for the Experimental Study of Education, Vol. I, 17-44.

Herring, J. P. (24). "Herring Revision of the Binet-Simon Tests." *J. of Ed. Psych.,* 15, 172-179.
Inglis, A. (22). *Intelligence Quotient Values.* Yonkers.
Jordan, A. M. (23). "The Validation of Intelligence Tests." *J. of Ed. Psych.,* 14, 348-666, and 414-428.
Kelley, T. L. (23). *Statistical Method.* New York.
Kelley, T. L. (27). *Interpretation of Educational Measurements.* Yonkers.
Kuhlmann, F. (27). "A Median Mental Age Method of Weighting and Scaling Mental Tests." *J. of Appl. Psych.,* 11, 181-198.
Kuhlmann, F. (28). "The Kuhlmann-Anderson Intelligence Tests Compared with Seven Others." *J. of Appl. Psych.,* 12, 545-594.
Lincoln, E. A. (28). "Studies of the Validity of the Dearborn General Intelligence Examinations." *J. of Ed. Psych.,* 19, 346-349.
Liu, H. C. (22). *Non-Verbal Intelligence Tests for Use in China.* T. C. Contribs. to Educ., No. 126, New York.
McAnulty, E. A. (28). "A Comparison of the Terman, National and the Stanford Binet Tests." *Ed. Res. Bull.,* Los Angeles, 8, No. 2, pp. 5-7.
McCall, W. A. (22). *How to Measure in Education.* New York.
McCall, W. A. (23). *How to Experiment in Education.* New York.
McCall, W. A., et al. (26). "Construction of Multi-Mental Scale." T. C. *Record,* 27, 394-415.
McGraw, M. L. and Mangold, M. C. (29). "Group Intelligence Tests in the Primary Grades." *Cath. Univ. Amer. Educ. Res. Bull.,* 4, No. 2.
Mitchell, D. (24). "Psychological Examination of Pre-school Age Children." *Ped. Sem.,* 31, 108-146.
Pintner, R. (18). *The Mental Survey.* New York.
Pintner, R. (27). "The Pintner-Cunningham Primary Test." *J. of Ed. Psych.,* 18, 52-58.
Pintner, R. and Marshall, H. (21). "A Combined Mental Educational Survey." *J. of Ed. Psych.,* 12, 32-43.
Popenoe, H. (27). "A Report of Certain Significant Deficiencies of the Accomplishment Quotient." *J. of Ed. Res.,* 16, 40-47.
Ruch, G. M. and Stoddard, G. D. (27). Tests and Measurements in High School Instruction. Yonkers.
Symonds, P. M. (27). *Measurement in Secondary Education.* New York.
Terman, L. M., et al. (17). *The Stanford Revision and Extension of the Binet-Simon Scale for Measuring Intelligence.* Baltimore.
Toops, H. A. and Pintner, R. (18). "A Chart for the Determination of I. Q. Values." *J. of Delinq.,* 3, 272.
Viele, A. B. (24-5). "A Study of Four Primary Mental Tests." *Elem. Sch. J.,* Vol. 25, 675-681.
Wilson, J. H. (24). "Comparison of Certain Intelligence Scales." *Brit. J. of Psych.,* 15, 44-63.

Witty, P. A. and Taylor, J. F. (29). "Some Results of the Multi-Mental Test." *J. of Educ. Psych.*, 20, 299-302.
Woolley, H. T. (15). "A New Scale of Mental and Physical Measurements for Adolescents." *J. of Ed. Psych.*, 6, 521-550.
Yepsen, L. N. (22). "A New I.Q. Slide Rule." *Sch. and Soc.*, 15, 596.

PART II

THE METHODS

CHAPTER VI

THE SCALES

This chapter will give a brief account of the various scales used for individual testing, with comments upon each. It is impossible to give all the necessary directions for the use of each scale. Those who desire to use any scale in actual testing must go to the original source and follow explicitly the directions of the author. No attempt will be made to describe the innumerable intelligence tests that have not been combined in some way or other into a scale. By a scale we mean a group of tests used in some special manner for the purpose of obtaining an intelligence rating, generally a mental age. By means of a single test we do not attempt to arrive at a mental age, although many single tests are of great help and form useful additions to the ordinary equipment of the psychologist. For a description of such tests, the reader must be referred to Whipple (14-15) and to Bronner et al (27). The purpose here, however, is to describe what we have designated "scales" in contradistinction to tests.

I. THE BINET SCALE

(a) *The 1905 Scale.*—The first appearance of the tests, familiar to us now as the Binet tests, was in the series of tests published in 1905, which we have called the 1905 Scale. Strictly speaking, it is not a scale in the later sense of that term. The tests are not grouped into age groups, but they are arranged in order of increasing difficulty. Below is printed a short description of the tests:

1. "Le Regard." Ability to follow with the eyes a moving object. A lighted match is slowly moved before the eyes to see if the child can follow the movement.
2. Prehension Provoked by a Tactile Stimulus. A piece of wood is brought into contact with the palm or back of the hand to see if the child will seize it without letting it fall and carry it to his mouth.
3. Prehension provoked by a Visual Stimulus. Cube of wood shown merely.
4. Recognition of Food. Present a piece of chocolate and a piece of wood.
5. Quest of Food Complicated by a Slight Mechanical Difficulty. Candy wrapped up in paper.
6. Execution of Simple Commands and Imitation of Simple Gestures.
7. Verbal Knowledge of Objects. Parts of the body and familiar objects.
8. Verbal Knowledge of Pictures. Ask where the things are on the picture.
9. Naming of Designated Objects. Common Objects on a picture must be named by the child.
10. Comparison of Two Lines. The child must tell which is the longer.
11. Repetition of Three Digits.
12. Comparison of Two Weights. Compare 3 and 12 grams.
13. Suggestibility. Asking for objects that are not present; for things on the picture by a nonsense word; comparison of lines, three unequal and last three equal.
14. Definition of Objects.
15. Repetition of Sentences.
16. Comparison of Two Objects. The difference between a fly and a butterfly, etc.
17. Memory for Things in a Picture. Thirteen common objects shown in a picture for thirty seconds, after which child repeats as many as he can.
18. Drawing a Design from Memory. Two designs are shown for 10 seconds, after which the child tries to draw them.
19. Repetition of Digits.

20. Resemblance of Known Objects. In what way are a poppy and blood alike? etc.
21. Comparison of Lengths of Lines.
22. Comparison of Weights.
23. Memory for Weights. After the blocks have been correctly placed in order of weight in Test 22, one is taken away and the subject must find out where the gap is.
24. Rhymes. Find rhymes to a given word.
25. Completion of Sentences. Find the right word to complete a sentence.
26. Makes up a Sentence including Three Given Words.
27. Comprehension of Questions. There are 25 questions from very easy to very hard.
28. Reversal of the Hands of a Clock. To be done from memory.
29. Paper Cutting. Paper folded twice and triangular piece cut out. Subject must draw result when paper is opened out.
30. Definitions of Abstract Terms.

Although Binet did not divide these tests into age groups, he gave general indications as to how far normal children of certain ages should go in the tests, e.g., age three up to about Test 9; age five up to about Test 14. Beyond this he suggests the range in different types of tests for various ages. In like manner he suggests only roughly the diagnosis of various degrees of feeblemindedness by means of these tests.

(b) *The 1908 Scale.*—Here the improvement over the previous scale lies in the fact that the tests are now grouped into years. It is the first age-scale.

Age III

1. Pointing to Nose, Eyes and Mouth.
2. Repetition of Short Sentences.
3. Repetition of Two Digits.
4. Enumeration of Objects in Pictures. The pictures are the Peddler and Boy; The Old Man and Woman on a Bench; The Man Standing on a Couch Looking Out of a Window.
5. Knows his Last Name.

Age IV

1. Knows Sex.
2. Names Familiar Objects.
3. Repetition of Three Digits.
4. Knows Longer of Two Lines.

Age V

1. Compares Two Boxes of Different Weight.
2. Copies Square.
3. Game of Patience. Rectangular card cut diagonally to be reconstructed according to a similar uncut card placed before the child.
4. Counts Four Coins.
5. Repeats a Sentence of Ten Syllables.

Age VI

1. Knows Right and Left.
2. Repetition of a Sentence of Sixteen Syllables.
3. Esthetic Comparison. Choose the prettier of three pairs of faces.
4. Definition of Familiar Objects.
5. Executes Three Commissions.
6. Knows Age.
7. Distinction between Morning and Afternoon.

Age VII

1. Unfinished Pictures. Tells what is missing.
2. Number of Fingers. Knows number on each hand and on both hands without counting.
3. Copy of a Written Model.
4. Copies a Diamond.
5. Repetition of Five Digits.
6. Description of Pictures.
7. Counts 13 Coins.
8. Knows Names of Four Common Coins.

Age VIII

1. Reads a Passage and Remembers Two Items.
2. Counts Nine Coins.
3. Names Four Colors. The four primary colors.
4. Counts Backwards from 20 to 0.
5. Writes from Dictation.
6. Comparison of Two Objects. Differences.

Age IX

1. Knows Date. Day of week, day of month, month and year.
2. Repeats Days of Week.
3. Makes Change.
4. Definitions Superior to Use.
5. Reads a Passage and Remembers Six Items.
6. Arranges Five Boxes in Order of Weight.

Age X

1. Repeats the Months of the Year.
2. Knows the Names of Nine Pieces of Money.
3. Uses Three Words in One Sentence.
4. Comprehension of Common-Sense Questions (Easy).
5. Comprehension of Common-Sense Questions (Difficult).

This last test Binet calls a transitional test between ages ten and eleven.

Age XI

1. Criticizes Absurd Statements.
2. Uses Three Words in One Sentence.
3. Names Sixty Words in Three Minutes.
4. Defines Abstract Words.
5. Disarranged Words Arranged into a Sentence.

Age XII

1. Repetition of Seven Digits.
2. Finds Three Rhymes for a Given Word.

3. Repetition of a Sentence of 26 Syllables.
4. Explains Problem Questions.
5. Interprets Pictures.

Age XIII

1. Paper Cutting. Described in 1905 Scale, Test 29.
2. Reversed Triangle. Rearranges two triangles in imagination and draws results.
3. Differences between Pairs of Abstract Terms.

The 1908 Scale divides the tests into age groups and this is the most significant advance over the 1905 Scale. In addition the number of tests has now increased from thirty to fifty-nine for ages three to thirteen. These do not include the first six tests of the 1905 Scale suitable for infants. In addition a few of the others have been dropped, e.g., Tests 9, 17, 20, 21, 23, 28. The method of giving several of the tests has been radically changed, and a great many new tests have been introduced. These changes and additions represent the result of the practical application of the 1905 Scale to many children. It is interesting to note the introduction of several tests depending more directly upon school knowledge, although Binet was evidently aware of the undesirability of that type of test. The following footnote of his is interesting: "These tests are not the first ones of which we had thought; if we keep them, it is after long trial; they appear to us all good and practical. But we are far from claiming that they are the best. Those who will take up this work after us will find better; they will certainly succeed in eliminating more strictly than we have been able to do, the tests that are influenced by education." (Binet and Simon, 16, Kite translation).

The 1908 Scale is also important, because in reference to it, Binet explained the method of calculating mental age for the first time. Briefly, a child who passes all or all but one of the tests of any age group is credited with that mental age to begin with (now commonly called the basal age). To this age is added as many years or fractions of years as he is en-

titled to for tests passed above his basal age, five or six additional tests being counted roughly as a year. This method of calculation is suggested by Binet and is used only in an approximate manner, without any of the niceties of present-day calculation. As a matter of fact he designates his cases as being so many years advanced or retarded and his smallest unit is a half year.

It was this 1908 Scale that was first introduced into this country and all the early work was done on this scale or on modifications of it. We shall, however, neglect these American modifications in the meantime and discuss Binet's final Revision of 1911, because all the useful American revisions at present have taken into account the changes made by Binet in the 1911 Scale.

(c) *The 1911 Scale.*—This was the result of two or three years' work on the 1908 Scale and the slight changes made arose from the criticisms of Binet and his co-workers. Binet tells us that he has omitted several tests which depend upon the scholastic ability of the child, such as reading and writing; also certain tests which are matters of knowledge dependent upon the environment, such as age, days of the week. In addition some tests were found too hard for the age at which they were placed and these have been moved, and a few new tests have been introduced. All tests for Age XI have been omitted and most of them moved to Age XII. Age XII and Age XIII become in effect Age XV and Adult. We also note that there are now five tests at each age, except at Age IV, where there are only four. Ages III, IV and V remain the same as in the 1908 Scale. For detailed description of tests, see Binet, A., and Simon, T., translations by Town and Kite respectively.

Age VI

1. Distinguishes Morning and Afternoon.
2. Defines Words in Terms of Use.
3. Copies a Diamond.

4. Counts 13 coins.
5. Esthetic Comparison.

Age VII

1. Knows Right and Left.
2. Description of Pictures.
3. Executes Three Commissions.
4. Gives Value of 3 Single and 3 Double Sous.
5. Names Four Colors.

Age VIII

1. Comparison of Two Objects. Differences.
2. Counts Backwards from 20 to 0.
3. Unfinished Pictures.
4. Knows Date.
5. Repetition of Five Digits.

Age IX

1. Makes Change.
2. Definitions Superior to Use.
3. Knows Value of 9 Pieces of Money.
4. Repeats the Months of the Year.
5. Comprehension of Easy Common-Sense Questions.

Age X

1. Arranges Five Boxes in Order of Weight.
2. Copies Two Designs from Memory.
3. Criticizes Absurd Statements.
4. Comprehension of Difficult Common-Sense Questions.
5. Uses Three Words in Two Sentences.

Age XII

1. Resists Suggestion (length of lines).
2. Uses Three Words in One Sentence.
3. Names Sixty Words in Three Minutes.
4. Defines Abstract Words.
5. Disarranged Words Arranged into a Sentence.

Age XV

1. Repetition of Seven Digits.
2. Finds Three Rhymes for a Given Word.
3. Repetition of a Sentence of 26 Syllables.
4. Interprets Pictures.
5. Explains Problem Questions.

Adult

1. Paper Cutting.
2. Reversed Triangle.
3. Three Differences between a President and a King.
4. Differences between Pairs of Abstract Terms.
5. Gives Sense of a Passage from Hervieu.

These three scales show the progressive development of the idea of measuring intelligence by age steps. From the first series of tests arranged in order of difficulty we proceed to the arrangement of tests according to age and then to the better standardization of each test according to age and to a uniform number of tests for each age. The tests themselves do not change materially from scale to scale. The conglomeration of tests has justified itself from the beginning as a reliable instrument for measurement, and in the successive revisions we see Binet at work adjusting more nicely his measuring rod. Had Binet lived longer, he would undoubtedly have continued adjusting it in order to measure still more accurately.

II. MODIFICATIONS OF THE BINET

As soon as Binet's work became known numerous investigators started to try out his method and naturally enough began to suggest various changes. It would be futile and unprofitable to attempt to record all the suggested changes. Many were stupid and made by individuals, who had merely read Binet's articles or, still worse, meager translations or accounts of the scale, and, without experimenting themselves,

felt called upon to pronounce judgment and to dictate how things should be done. From serious workers came many valuable criticisms which were all more or less taken account of by the chief revisions that were made later. We shall describe here a few of the more important revisions and modifications of the Binet, which have been used in America.

(a) *The Goddard Revision.*—Goddard (11) was the first to introduce the Binet Scale into America and for a long time his Revision, first published in 1910, was the one most commonly used and the best standardized. In his standardization on American children he found many tests misplaced in the French Scale and, therefore, some change occurs in the position of some of the tests. He, of course, made all the necessary changes in order to adapt the tests to American conditions. He chose different pictures from those used by Binet, and introduced a few new tests.

The importance of the Goddard Revision is due to the fact that it was the first well-standardized scale for American children, and from his work with the feebleminded he gave us tentative diagnostic limits on his scale. These limits at first were that two years' retardation was probably indicative of feeblemindedness. Later on these limits were revised to read two years' retardation below age nine and three years' above that age. In calculating mental age the method of Binet was followed, but it was more rigidly carried out and fractions of a year were now used. The method employed by Goddard was to allow a year for any five tests passed above the basal year as Binet had done, and in addition any extra tests were indicated by an exponent. Thus a child having passed all the 8-year tests and also three 9-year tests would be given a mental age of 8^3, the exponent 3 indicating three-fifths of a year. Little use was made of the XV and Adult tests, and the standardization at these ages was very unsatisfactory.

(b) *The Kuhlmann Revision.*—Two revisions of the Binet Scale have been published by Kuhlmann (12 and 22). The

first one adhered closely to the original Binet Scale, while the second one, published ten years later, introduced a great number of new tests. The general nature of the 1922 Kuhlmann Revision may be described as follows: elimination of nineteen of the original tests because they were found to be unsatisfactory; increase of the number of tests to eight in each age-group above two years; extension of the original scale at both upper and lower ends. There are 129 tests in place of the 56 tests of the original scale.

The most original contribution of Kuhlmann is the standardization of tests for children below the age of three. Some such tests were proposed by Binet in the 1905 Scale, but they were few in number and unstandardized. We give below Kuhlmann's tests for children below the age of three:

Age Three Months

1. Carrying hand or object to mouth.
2. Reactions to sudden sound. (Clap hands, etc.).
3. Binocular co-ordination. (Moving candle.)
4. Turning eyes to object in marginal field of vision.
5. Winking at an object threatening the eyes.

Age Six Months

1. Balancing head. Sits upright.
2. Turning the head towards a source of sound.
3. Opposing thumb in grasping.
4. Prolonged holding of object placed in hand.
5. Reaching for seen objects.

Age I

1. Sitting and standing.
2. Speech (Imitates ba, dada, mama, etc.).
3. Imitation of movements. (Shakes rattle, rings bell, etc.)
4. Marking with pencil. (Imitation.)
5. Recognition of objects. (Shows preference.)

Age Eighteen Months

1. Drinking from a glass. (Several swallows.)
2. Feeding with spoon or fork.
3. Speech (mama, papa, baby, yes, no, cat, man).
4. Spitting out solids.
5. Recognition of objects in pictures.

Age II

1. Pointing out objects in pictures.
2. Imitation of simple movements. (Raise arms, clap hands.)
3. Obeying simple commands. (Catch ball, throw ball, etc.)
4. Copying a circle. (Any rough circular movement.)
5. Removal of wrapping from food before eating.

The mental age on the Kuhlmann is calculated in the usual way. In all ages above Age II there are eight tests. Each test passed, therefore, counts as one eighth of a year. A useful table for converting mental age as expressed in years and eighths into I.Qs. is given by the author. Kuhlmann's general interpretation of the significance of the various I.Qs. is as follows:

Grade Terms	*Range in I.Q.*
Idiots	0-24
Imbeciles	25-49
Morons	50-74
Borderline	75-84
Dull	85-94
Average	95-104
Bright	105-114
Very Bright	115-124
Superior	125-149
Very Superior	150-174
Precocious	175 and over

(c) *The Point Scale.*—A Point Scale for measuring mental ability appeared first by Yerkes (15), Bridges and Hardwick.

A later revision was published by Yerkes (23) and Foster. This later revision offers three scales:—the pre-adolescent scale; the adolescent and adult scale; the infant scale. Only the first of these has been satisfactorily completed and we shall confine our description to it. Although making use of the original Binet tests, the Point Scale is avowedly different in method.

The authors of the Point Scale seem to have objected to two things in the original Binet system, namely, the grouping of tests into age-groups and the system of scoring a test either right or wrong without allowing any partial credit. The latter system they call the "all or none" method of giving credit, and in its place they have developed a system of partial credits according to the number of items in a test that are answered correctly. The result of these two changes is that the characteristic grouping of tests of the Binet Scale into different age groups now disappears, and in its place appears a list of tests, all of which are to be given to each child examined, and for each of which a certain amount of credit or a certain number of points is allotted. The total number of points achieved by any individual gives the intelligence rating of that individual.

The point system, i.e., the system of giving varying amounts of credit according to the quality of the performance, is unquestionably very useful and perhaps fundamentally the ideal system in mental testing. It forms the basis of almost all our group methods of mental testing. Ideally, however, it presupposes a test in which the poorest can score just a little and the very best not quite the maximum. Many group tests fulfill this condition for a certain definite range of ages. The application of this system to tests of the Binet type results merely in a system of partial credits, for very few of the Binet tests cover a wide range of ages. Tests such as choosing the prettier of two pictures, or detecting missing parts are obviously suitable for young children; the tests of dissected sentences or of drawing designs from memory are limited to

older children. A point system of rating demands, however, that we give every child a trial on all the tests, because his score represents his total achievement on the whole scale. The consequence is that a point scale, using tests of the Binet type, involves much useless testing. For such tests the original Binet system of age-groups is unquestionably the best, unless we adopt a modified point system such as the one constructed by Herring to be described later.

The Point Scale consists of twenty tests, beginning with very easy tests, such as "choosing the prettier of two pictures," and continues to harder tests, such as definitions and analogies. All of the tests, except the analogies, are modifications of the original Binet tests. After each test has been given the credits obtained by the subject are noted in a special column of the test blank. At the end of the examination the total credits obtained forms the basis of the mental rating. The maximum possible is 100.

By means of a table, the equivalent mental age for any specific score can be ascertained. These mental ages range from 4 to 18. Then the I.Q. may be calculated in the usual way.

With the original Point Scale, Yerkes recommended the use of the coefficient of Intellectual Ability (C.I.A.). This coefficient is arrived at by dividing the score obtained by the standard score for the chronological age of the subject examined. Thus, a score of 20 made by a five year old would give a C.I.A. of 20/22 or .91. The same score made by a six year old would give a C.I.A. of 20/28 or .72, and so on. These coefficients and the mental age equivalents of any score are most rapidly and accurately obtained from a chart such as the one published by Pintner and Toops (17).

The standardization of the scale is good. It was in fact so superior to the standardization of most scales in use at the time of the appearance of the Point Scale, that from that time on, until the appearance of the Stanford Revision, the Point Scale was used by many psychologists. The superior stand-

ardization of the Stanford Revision and its wider range of applicability has resulted in little use of the Point Scale at the present time. We must not forget, however, that the Point Scale has proved of decided value, and that it stimulated thought and experimentation on the method of scale construction.

(d) *The Stanford Revision.*—The Stanford Revision of the Binet Scale by Terman and his co-workers is, as its name indicates, an extension, elaboration and thorough revision of the original Binet Scale. A guide for the use of the Scale appeared in 1916 (Terman, 16), and an account of its construction appeared in 1917 (Terman, 17), but historically we note its beginning in 1910 when Terman (12) and Childs began their critical evaluation and tentative revision of the Binet 1908 Scale. From 1910 onwards till the year of publication mentioned above, Terman and his co-workers were at work upon the scale and ultimately produced the Stanford Revision. No other scale has had such a thorough and extensive foundation.

The Binet tests of both the 1908 and 1911 scales form the foundation of the work and the Binet methods of age-grouping and calculation of mental age have been retained. The original Binet Tests have, however, been made much clearer and many ambiguities in procedure eliminated. Many new tests have been introduced, "counting both regular and alternative tests, the revision contains 90 tests, as contrasted with 54 in the Binet 1911 Series" (Terman, 17). The placing of tests at the proper age has been more carefully worked out than in the original Binet or in any of the other revisions. The device of having six tests in each age group has made the calculation of mental age at once simpler and also more accurate. Furthermore, a very definite attempt was made to standardize the scale as a whole so that the average child at each age would test exactly at age. This lack of adjustment of the scale as a whole was very noticeable in the original Binet and in the Goddard Revision, so much so that an average

three year old tested more than three mentally and an average eleven year old less than eleven.

The data used in the Stanford Revision include 2,060 examinations. All tests before being admitted into the scale were carefully scrutinized as to their validity. Terman (17) says, "there are three criteria which a test must satisfy before it can be accepted as a valid measure of intelligence."

1. Age Increase. A test must show an increase in the percentage of children that pass it from year to year.
2. Coherency. A test must be coherent with the scale as a whole; it must be measured up against the entire scale. This is shown by the percentages of dull, normal and bright children, as diagnosed by the whole scale, that pass each particular test. It is a question of the extent of the correlation between each test and the whole scale.
3. World Success. A test, or rather a scale of tests, must show a certain amount of agreement with the rating of an individual's intelligence, such as is reflected by his success in the world, by his success in school, by judgments of his intelligence by teachers and friends.

These are three criteria mentioned by Terman, and we have discussed them or similar ones more fully in a previous chapter. They are worthy of mention here, however, so that we may fully appreciate the conscious application of principles underlying the Stanford Revision.

A summary of the tests in their age grouping is as follows: *

Year III

1.* Points to parts of body—nose, eyes, mouth, hair.
2.* Names familiar objects—key, penny, knife, watch, pencil.
3.* Pictures—enumeration.
4. Gives sex.
5. Gives last name.

6.* Repeats 6-7 syllables—(*a*) I have a little dog. (*b*) The dog runs after the cat. (*c*) In summer the sun is hot.

Year IV

1.* Compare lines.
2. Discrimination of forms.
3.* Counts four pennies.
4.* Copies square.
5.* Comprehension. What must you do (*a*) when you are sleepy; (*b*) when you are cold; (*c*) when you are hungry.
6. Repeats four digits. 4-7-3-9; 2-8-5-4; 7-2-6-1.

Year V

1.* Comparison of weights, 3 and 15 grms.
2.* Colors. Red, yellow, blue, green.
3.* Esthetic comparison.
4. Definitions, use or better. Chair, horse, fork, doll, pencil, table.
5. Patience, or divided triangle.
6.* Three commissions. Puts key on chair, brings box, shuts door.

Year VI

1.* Right and left. R. hand; L. ear; R. eye.
2.* Mutilated pictures.
3.* Counts 13 pennies.
4.* Comprehension. What's the thing to do:
   (*a*) If it is raining when you start to school?
   (*b*) If you find that your house is on fire?
   (*c*) If you are going some place and miss your car?
5. Coins. Nickel, Penny, Quarter, Dime.
6. Repeats 16-18 syllables.
   (*a*) We are having a fine time. We found a little mouse in the trap.
   (*b*) Walter had a fine time on his vacation. He went fishing every day.
   (*c*) We will go out for a long walk. Please give me my pretty straw hat.

Year VII

1.* Fingers. Number on R. hand; L. hand; both.
2.* Pictures, description or better.
3. Repeats 5 digits. 3-1-7-5-9; 4-2-8-3-5; 9-8-1-7-6.
4. Ties bow knot. Model shown. Single bow half credit.
5.* Gives differences. (*a*) fly and butterfly, (*b*) stone and egg, (*c*) wood and glass.
6.* Copies diamond.

Year VIII

1. Ball and field. (Inferior plan or better.)
2.* Counts backward 20 to zero.
3.* Comprehension. What's the thing for you to do:
   (*a*) When you have broken something which belongs to someone else?
   (*b*) When you are on your way to school and notice that you are in danger of being tardy?
   (*c*) If a playmate hits you without meaning to do it?
4.* Similarities. (*a*) Wood and coal; (*b*) apple and peach; (*c*) iron and silver; (*d*) ship and automobile.
5. Definitions superior to use. (*a*) Balloon; (*b*) Tiger; (*c*) Football; (*d*) Soldier.
6.* Vocabulary, 20 words.

Year IX

1.* Date: week, month, day of month, year.
2.* Weights.
3. Makes change. 10 − 4; 15 − 12; 25 − 4.
4.* Repeats 4 digits backwards. 6-5-2-8; 4-9-3-7; 8-6-2-9.
5.* Three words in a sentence. (*a*) Boy, river, ball; (*b*) work, money, men; (*c*) desert, rivers, lakes.
6. Rhymes. Three rhymes for each word. (*a*) day; (*b*) mill; (*c*) spring.

Year X

1.* Vocabulary, 30 words.
2.* Absurdities.
3. Designs.
4. Reading and report.

New York. | September 5th. | A fire | last night | burned | three houses | near the center | of the city. | It took some time | to put it out. | The loss | was fifty thousand dollars, | and seventeen families | lost their homes. | In saving | a girl | who was asleep | in bed, | a fireman | was burned | on the hands.

5.* Comprehension.

(*a*) What ought you to say when someone asks your opinion about a person you don't know very well?

(*b*) What ought you to do before undertaking (beginning) something very important?

(*c*) Why should we judge a person more by his actions than by his words?

6.* Sixty words in three minutes.

Year XII

1.* Vocabulary, 40 words.
2. Abstract words. (*a*) Pity, (*b*) Revenge, (*c*) Charity, (*d*) Envy, (*e*) Justice.
3. Ball and field (superior plan).
4.* Dissected sentences.

(*a*) for the started an we country early at hour.

(*b*) to asked paper my teacher correct I my.

(*c*) a defends dog good his bravely master.

5.* Fables.
6.* Repeats 5 digits backwards. 3-1-8-7-9; 6-9-4-8-2; 5-2-9-6-1.
7.* Pictures, interpretation.
8.* Gives similarities (*a*) snake, cow, sparrow. (*b*) book, teacher, newspaper. (*c*) wool, cotton, leather. (*d*) knife-blade, penny, piece of wire. (*e*) rose, potato, tree.

Year XIV

1.* Vocabulary, 50 words.
2. Induction test. (Gets rule by 6th folding.)
3.* President and king. Difference between.
4.* Problems of fact.

5.* Arithmetical reasoning.
6. Clock. Reverse hands; (*a*) 6: 22, (*b*) 8: 10, (*c*) 2: 46.

Year XVI

1.* Vocabulary, 65 words.
2.* Fables.
3. Difference between abstract words. (*a*) laziness and idleness, (*b*) evolution and revolution, (*c*) poverty and misery, (*d*) character and reputation.
4.* Problem of the enclosed boxes. One large box containing (*a*) 2 smaller, 1 inside of each; (*b*) 2 smaller, 2 inside of each; (*c*) 3 smaller, 3 inside of each; (*d*) 4 smaller, 4 inside of each.
5.* Repeats 6 digits backwards. 4-7-1-9-5-2; 5-8-3-2-9-4; 7-5-2-6-3-8.
6. Code, writes "come quickly."

Year XVIII

1.* Vocabulary, 75 words.
2. Paper cutting test. Draws folds and locates holes.
3.* Repeats 8 digits forwards. 7-2-5-3-4-8-9-6; 4-9-8-5-3-7-6-2; 8-3-7-9-5-4-8-2.
4.* Repeats thought of passage heard.
5.* Repeats 7 digits backwards. 4-1-6-2-5-9-3; 3-8-2-6-4-7-5; 9-4-5-2-8-3-7.
6. Ingenuity test.
   (*a*) A mother sent her boy to the river to get seven pints of water. She gave him a 3-pint vessel and a 5-pint vessel. Show me how the boy can measure out exactly 7 pints without guessing at the amount. Begin by filling the 5-pint vessel.
   (*b*) Same, except 5 and 7 given to get 8. (Begin with 5.)
   (*c*) Same, except 4 and 9 given to get 7. (Begin with 4.)

The fables used in Ages XII and XVI are entitled (*a*) Hercules and the Wagoner; (*b*) the Milkmaid and her Plans;

(*c*) The Fox and the Crow; (*d*) The Farmer and the Stork; (*e*) The Miller, his Son, and the Donkey. A score of two is given for a generalized interpretation, and a score of one for an interpretation that does not go beyond the characters in the fable.

The words used in the Vocabulary Test consist of two series of fifty words each. The idea of the test is to find out whether the child knows the words and not whether the child can define them accurately. A word is considered correct if the subject shows he knows any meaning of the word, however poorly he may define it. The arrangement of the words in the list is from very easy to very hard.

It will be noticed at once from this summary of the Stanford Revision how greatly the scope of the original Binet Scale has been extended and how the whole scheme has been made more consistent and uniform.

The Short Scale.—Because the use of the complete scale takes a fairly long time, and because many occasions arise when it is necessary to get an estimate of a child's intelligence as quickly as possible, Terman has indicated a certain number of tests at each age which may be used as a short scale. The correlation between the short and the long scales is high. In spite of this, however, the long scale will give the most accurate measure of the individual's mentality and should always be used when possible. The tests to be used in giving the short scale are indicated by an asterisk after the test number.

The Mental Age.—It will be noted that there are six tests at each age from III to X inclusive, and, therefore, the value of each test at these ages is two months. After age X the difference between intelligence from one age to the next becomes gradually less and less, and it becomes harder to find appropriate tests. Terman has, therefore, skipped one year in each case and standardized tests for ages XII, XIV, XVI and XVIII. Furthermore, the number of tests at age XII is eight. Since the tests for age XII cover the ages from X to XII, each test is equivalent to 3 months ($3 \times 8 = 24$ mos.). For

the other ages, Terman gives the following equivalents in months: age XIV — 4 months each; age XVI — 5 months each; age XVIII — 6 months each. It will be noted that by counting the actual number of months if all tests are passed, the total is 19 years, 6 months. This is the highest mental age that can be obtained on the Stanford Revision. If the Short Scale is used the tests are of course given higher values, e.g., each test in ages III to X counts 3 months, and so forth.

The method of calculating mental age is the same as suggested by Binet, namely the addition to the basal age of all tests passed above the basal age. The advantage of having 6 tests at each age, each equivalent to two months, is at once apparent in contrast to the old Binet Scale with its five tests necessitating the calculation of tenths of a year. The mental age on the Stanford is easily obtained by adding to the basal age the number of months' credit obtained by passing tests at higher ages. The result is, therefore, always stated in years and months. It is customary to use a dash between the years and months, e.g., 10 – 6, meaning thereby 10 years, 6 months, and one must not read this as ten years and six tenths.

Age XVI is called by Terman, Average Adult Intelligence, because it seemed to coincide with the ability of average adults as far as he could estimate at that time. From the results of the mental testing in the U. S. Army during 1917-1919, it would seem that average intelligence is more nearly somewhere between a mental age of 13 and 14 on the Stanford Scale. At the present time, therefore, it is well to regard Age XIV as average adult, Age XVI as Superior Adult, and Age XVIII as Very Superior Adult.

The Intelligence Quotient.—The I.Q. is the mental age divided by the chronological. Both ages, chronological and mental, should be stated in terms of years and months in order to get as accurate an I.Q. as possible. If, as is usual the ages are written with a dash between years and months, the examiner must be careful to note that this is not a decimal point, and, therefore, the two values cannot be divided until

they have both been converted into months. For example, Chron. Age = 10 - 8; M.A. = 8 - 4.

$$\text{I.Q.} = \frac{\text{M.A.}}{\text{C.A.}} = \frac{8-4}{10-8} = \frac{100 \text{ months}}{128 \text{ months}} \times 100 = 78$$

This rule for the calculation of the I.Q. holds good for all ages up to the Adult. All adults, regardless of age, will of course be considered as of that chronological age decided upon as equivalent to average adult intelligence. Terman gives as his rule the use of age 16 as the divisor for all adults, i.e., for all subjects who are 16 or over. In view of what has been said above as to the average age of adult intelligence, the present writer uses and recommends for use age 14 as the divisor for the calculation of the I.Q. for all adults, i.e., for all subjects who are 14 or over. The use of 14 rather than 16 agrees well with more recent considerations as to the upper mental age of feebleminded adults. It is, of course, incorrect to assume any limit of mental age above which no person is to be considered feebleminded, for we must remember that feeblemindedness is a social as well as a psychological concept. If, however, we consider the psychological criterion of feeblemindedness apart from any other criteria, we are rapidly coming to the opinion that no adult of mental age 10 and above is to be considered feebleminded. This being so, and regarding an I.Q. of 70 as the upper limit of feeblemindedness, we note that age 14 is the most suitable age to use in the calculation of the I.Q. of all individuals age 14 and above, because this makes a mental age of about 9 - 10 the upper limit of feeblemindedness. An adult testing M.A. 10, will have an I.Q. (using age 14) of 71, but if we use 16, the I.Q. will be 63, i.e., definitely feebleminded. Again using 16 as our divisor, and I.Q. below 70 as presumably feebleminded, we shall consider feebleminded all adults who test up to a mental age of 11 - 1 because M.A. 11 - 1 divided by age 16 gives an I.Q. of 69. If we use as our divisor age 14, all adults testing between M.A. 10 and M.A. 11 - 1 will have I.Qs. ranging from 70 to 79 and, therefore, will not be

considered feebleminded, as far as the psychological criterion is concerned. From these considerations it is much safer at the present time to consider 14 as the average mental age of adults and use it in the calculation of the I.Q. of all individuals who are chronologically 14 or above.

(e) *The Herring Revision.*—The latest revision of the Binet Scale has been constructed by Herring (22). It is an attempt on the part of the author to construct a scale having the same meaning and significance as the Stanford Revision. The correlation between the two scales for 126 children ranging in age from 8 to 13 is .98. Another 154 cases aged 4 to 18 also gave a correlation of .98. To all intents and purposes we would seem to have here an alternative form of the Stanford Revision. As the Herring Scale is made up of similar but different material, it is of great use in the retesting of children. It will eliminate such practice effect as always results from giving the same scale a second time.

The Herring Scale is a point scale in as much as each test is allotted a number of points. A score in the examination as a whole is the sum of the scores obtained for the separate tests. In order to obviate the necessity for giving all the tests to every subject, the scale is divided up in five groups of tests, Groups A to E. The score made in the first group of tests, Group A, determines the tests which are to be given in the succeeding groups. This circumvents the objection raised against the Yerkes-Bridges Point Scale, with which scale, it was necessary to give all of the tests to each subject.

Among the new tests found in the Herring Revision and not found in the Stanford Revision are the following:—Number Series Completion; Size Comparisons; Proverbs; Following Directions; Generalization; News Route. The number of tests in which reading is required is much greater than in the Stanford. The illiterate child or non-English-speaking child would seemingly be even more severely handicapped than on the Stanford.

After the total score has been obtained by adding the scores

on all of the tests, the mental age is determined by the tables of norms furnished by the author. The I.Q. is then calculated in the usual manner. The standardization of the Herring Scale is not the same as the Stanford for the upper part of the distribution. This results in much lower I.Q.'s on the Herring for children of superior intelligence. For children with I.Q.'s on the Stanford above 130 an average drop of 17 points in I.Q. has been noted on the Herring Revision (Carroll and Hollingworth, 30).

III. PERFORMANCE SCALES

The Binet Scale and its modifications consist of a miscellaneous group of tests given by means of an oral interview between subject and examiner. Very few of the tests are of the performance type, where the subject is expected to do things rather than to talk about things. This limits the use of the Binet type of scale to subjects capable of understanding the English language and capable of responding in that language. It also limits these scales to the testing of verbal or abstract intelligence rather decidedly and gives little or no opportunity for testing concrete or mechanical intelligence. Those who believe that intelligence may show itself in reactions to concrete material, as well as in verbal reactions, have therefore stressed the value of performance tests. We shall describe some of the more commonly used performance scales, but we shall not attempt to survey all the single performance tests which are in use, but which have not been combined into scales. For such single performance tests the best source is Bronner (27).

(a) *The Pintner-Paterson Performance Scale.*—As its name implies this scale (Pintner, 17) is made up of tests calling for a motor rather than a verbal response. The subject is required to do something rather than to say something. All of the tests can be presented without language and no language is required

for the response. Hence the scale is admirably adapted for the testing of foreign and deaf subjects as well as the ordinary English-speaking hearing subject. It has proved a valuable supplement to scales of the Binet type.

The scale consists of 15 tests as follows:

1. Mare and Foal Board. This is a picture board of a mare and foal with a number of cut-outs which the subject has to put in the correct places. It is very simple and resembles a child's game and serves as a very good introduction for children. Time and number of errors are recorded.
2. Seguin Form Board. Ten blocks representing common geometrical forms are to be placed in their appropriate places. The time of the shortest of three trials is recorded.
3. Five Figure Board. Five geometrical figures each divided into two or three pieces are to be placed in their appropriate places. Time and number of errors are recorded.
4. Two Figure Board. Nine pieces to be placed in two spaces. Time and number of moves are recorded.
5. Casuist Board. A more difficult board, consisting of four spaces into which have to be fitted twelve blocks. Time and number of errors are recorded.
6. Triangle Test. Four triangular pieces to be fitted into the board. Time and errors are recorded.
7. Diagonal Test. Five variously shaped pieces are to be fitted into a rectangular frame. Time and errors are recorded.
8. Healy Puzzle A. Five rectangular pieces are to be fitted into a rectangular frame. Time and moves are recorded.
9. Manikin Test. Subject has to put together legs, arms, head and body to form a man. There is no board into which the pieces fit. Quality of performance is scored.
10. Feature Profile Test. In the same manner as in the previous test, subject has to put together pieces to form a head. Time is recorded.
11. Ship Test. This consists of the picture of a ship cut into ten pieces of the same size and shape which are to be fitted together properly in a rectangular frame. Quality of performance scored.

12. Picture Completion Test. Subject is required to select the appropriate block out of many possible blocks to complete the picture. Quality of performance scored.
13. Substitution Test. A sheet of paper with rows of geometrical figures upon which the subject has to write the proper digit following the key at the top of the page. Time and errors compounded into a score.
14. Adaptation Board. This is a simple test for measuring the ability of the subject to keep his attention upon a moving board. Number of correct moves is recorded.
15. Cube Test. Four cubes are tapped in a certain order and the subject is required to watch and then imitate the movement. Number of combinations correctly imitated is recorded.

Various methods for the calculation of mental age by means of these tests have been suggested by the authors. They show how they may be shaped into a year scale or a point scale. They also show how a mental age may be derived from tables of median performance and also how a percentile rating of the child may be obtained. Those who have made practical use of the scale find the median mental age method of computing mental age the simplest and most accurate. The various times, errors, moves, scores, etc., for the tests are each converted into a mental age by means of the table of norms, and the median of all these mental ages is the mental age of the child. By this method an approximate mental age can be obtained from any number of tests.

The Short Performance Scale.—Ten of the above tests have been chosen by Pintner and Spaid for a short performance scale and they are to be recommended for general testing purposes. The ten tests used are numbers 1, 2, 3, 4, 5, 9, 10, 11, 12, and 15 of the list given above and are shown in Figure 14.

Tests 9 and 10, the Manikin and Feature Profile, are used as one test because the norms for the former run from age 4 to 8 and for the latter from age 10 to 15. The tests are given and scored as in the long scale and the median mental age is

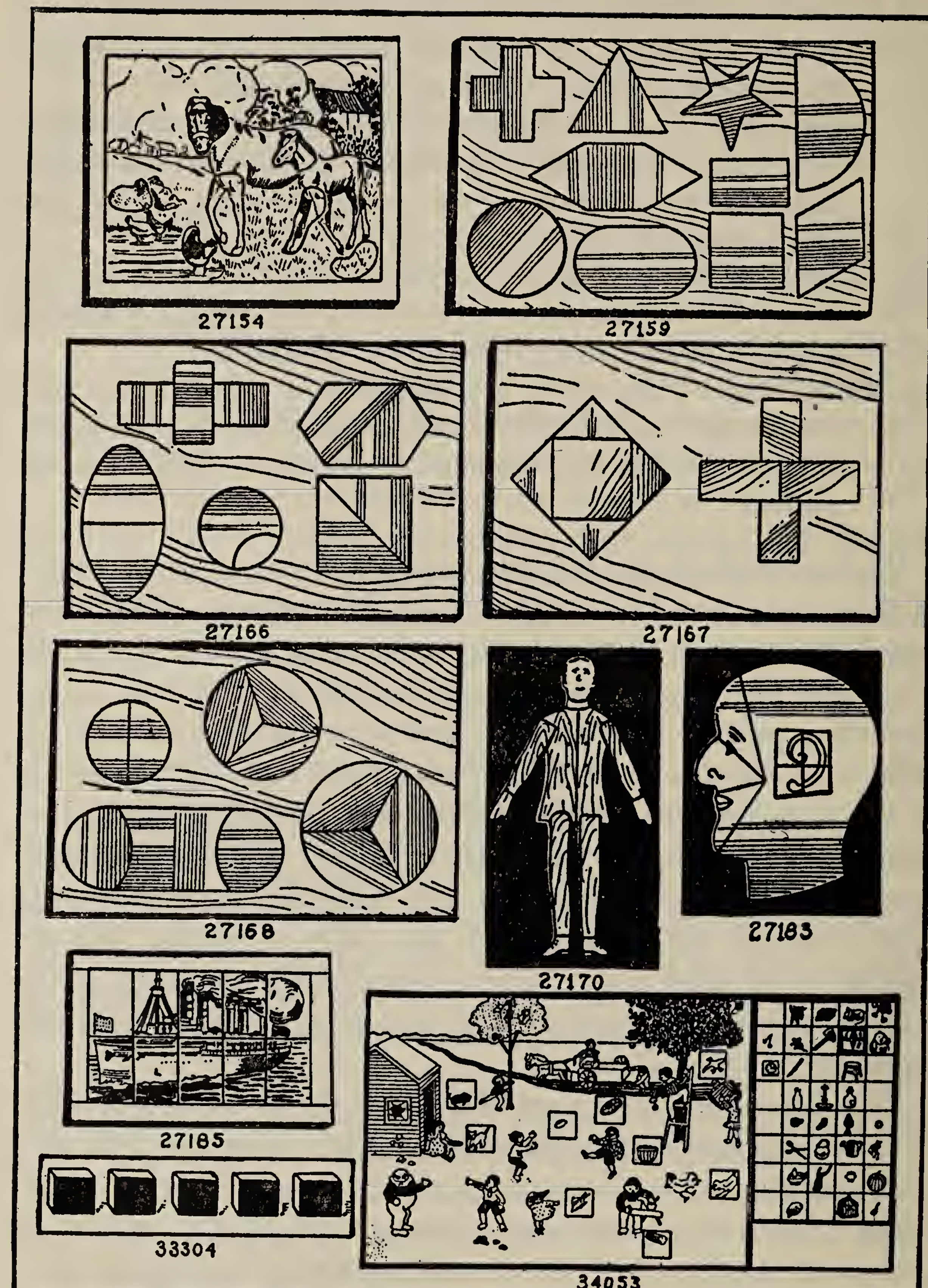


FIG. 14.—The Pintner-Paterson Short Performance Scale. (Courtesy of C. H. Stoelting Company, Chicago.)

used in the computation of mental age. A further convenience has been introduced by reducing the size of several of the form-boards so that all of the tests of the Short Scale can be conveniently carried in a small case. This makes the Performance Scale suitable for the psychologist who has to take his materials with him into the school or hospital.

(b) *The Army Performance Scale.*—Like the Pintner-Paterson Performance Scale, this scale (Memoirs, 21) is also designed for use without language. It was used in the testing of foreign and illiterate recruits in the army. There are ten tests, the first three of which are taken from the Pintner-Paterson Scale. The tests are: (1) Ship, (2) Manikin and Feature Profile, (3) Knox Cube, (4) Cube Construction, (5) Form Board (Dearborn's), (6) Copying Designs, (7) Digit-Symbol, (8) Maze (Porteus), (9) Picture Arrangement (Fraser and Whipple), (10) Picture Completion (Healy).

The tests are scored by assigning so many points to various types of performance and these scores are then converted into a weighted score for each test. The total weighted score represents the individual's performance on the whole scale. This total score can be converted into the letter ratings which were used in the army. This scale has not been standardized on children, but equivalent mental ages for various scores have been calculated. A short form of this scale consists of tests 1, 2, 3, 4, 6, 7, and 8.

(c) *The Drever-Collins Performance Scale.*—This scale has been constructed by Drever (28) and Collins primarily for the testing of the deaf, but it has been used with hearing children as well. It consists of the following tests: (1) Block Design Test (Kohs), (2) Cube Test (Knox and Pintner), (3) Domino Test (Drever and Collins), (4) Size and Weight Test, (5) Manikin and Profile Test (Pintner and Paterson), (6) Form-boards Test (Pintner's Two Figure Board and Healy's Puzzle A), (7) Cube Construction, (8) Picture Completion Boards (Drever and Collins, and Healy's Picture No. 1).

The tests are scored by allotting so many points to each test.

The maximum score is 148. There are tentative age norms for both hearing and deaf children.

(d) *Other Performance Scales.*—Arthur (28) has constructed a performance scale made up of eight of the Pintner-Paterson series together with the Kohs Block Design and Porteus Maze Tests. Points are allotted for all degrees of ability on each test. The total number of points can then be converted into a mental age.

Kent and Shakow (28) have combined a series of form-boards into a kind of performance scale, and Squires (26) has partially worked out a series of fourteen performance tests for higher levels of ability (about high school and college ability).

IV. TESTS FOR PRE-SCHOOL CHILDREN

(a) *The Merrill-Palmer Scale.*—This scale (Stutsman, 31) consists of seventy-two tests arranged in groups for children from 18 to 53 months. The tests are in groups suitable for each six-month interval. The names of some of the tests are:

Test 2—Throwing Ball	18 to 23 months
Test 9—Repetition of Words	18 to 23 months
Test 15—Wallin Peg Board A	24 to 29 months
Test 27—Three Cube Pyramid	30 to 35 months
Test 40—Little Pink Tower	36 to 41 months
Test 53—Decroly Matching Game	42 to 47 months
Test 64—Action Agent	48 to 53 months

For the most part the tests are of the performance type, but the directions are given in English, so that it is not a non-language scale. In addition there are a few verbal tests.

The total score on the tests can be interpreted in terms of mental age, standard deviation or percentile rank.

(b) *Gesell's Development Schedules.*—From an intensive study of many infants during the first three years of life, Gesell (25, 29) has arranged a series of items for various levels. Some of these items are intelligence tests in the usual mean-

ing of that word, but many are motor or language reactions. The use of these schedules by expert examiners gives a good measure of the total mental and physical development of the child. Accelerated or retarded development can be very definitely measured. The levels represented in the Gesell schedules are each month from 1 to 10 inclusive; then month 12, 15, 18, 21, 24 and 30. The reader will better appreciate what these levels are by the following samples taken from Gesell * (29):

Normative Summary for One Month Level:

Motor Development: (a) Lifts head from time to time when held to the shoulder; (b) Makes crawling movements when laid prone on flat surface; (c) Lifts head intermittently, though unsteadily, when in this prone position; (d) Turns head laterally when in prone position.

Language: (a) Gives definite heed to sound; (b) Has differential cries for discomfort, pain, and hunger.

Adaptive Behavior: (a) Stares at a window or at massive objects; (b) Gives visual heed to conspicuous moving objects; (c) Gives transient visual regard to the red ring; (d) Retains definite hold of the ring when it is placed in the hand.

Personal-Social Behavior: (a) Makes tactually perceptible postural adjustments when taken up by the examiner; (b) Shows selective regard for the face.

Normative Summary for Four Months Level:

Motor Development: (a) Holds head steady when carried or when swayed; (b) Lifts head and shoulders in dorsal position as an effort toward sitting; (c) Sits with resistant body posture when supported by pillows; (d) Hands no longer predominantly clenched, but frequently open.

* Gesell (29) specifically states that, although the items have been arranged in age groups, they are not to be simply scored plus and minus and then compounded into a mental age. The use of these normative summaries presupposes a large amount of clinical familiarity with infants of every age. All test materials referred to in these summaries are described by Gesell in *The Mental Growth of the Pre-school Child.* Quotations by permission of the Macmillan Company, publishers.

Language: (a) Laughs aloud; (b) Responds vocally when socially stimulated; (c) Vocalizes in self-initiated sound play.

Adaptive Behavior: (a) Closes in with both hands on dangling ring, when in dorsal position; (b) Manipulates table edge when held in lap; (c) Regards one-inch cube on table; (d) Turns head in pursuing slowly vanishing object.

Personal-Social Behavior: (a) Inspects own hand in play; (b) Plays in simple manner with rattle; (c) Splashes with hand in bath; (d) Makes definite anticipatory adjustment to being lifted.

Normative Summary for Eight Months Level:

Motor Development: (a) Sits momentarily without support; (b) Raises self to sitting position; (c) Picks up pellet with partial finger prehension.

Language: (a) Gives vocal expression to recognition; (b) Vocalizes in interjectional manner.

Adaptive Behavior: (a) Definitely looks for fallen spoon; (b) Utilizes handle in lifting inverted cup; (c) Shows manipulatory interest in details of bell.

Personal-Social Behavior: (a) Shows definite responsiveness to frolic play; (b) Pats or smiles at mirror image; (c) Restores bottle to mouth; (d) Shows interest in throwing and sound production play.

Normative Summary for Twelve Months Level:

Motor Development: (a) Walks with help; (b) Lowers self from standing to sitting position; (c) Holds crayon adaptively to make stroke.

Language: (a) Says two "words"; (b) Adjusts to simple verbal commissions; (c) Places cube in or over cup on command.

Adaptive Behavior: (a) Imitates scribble or rotary spoon rattle; (b) Adjusts round block to form-board or rod to hole; (c) Uses string adaptively to pull ring; (d) Secures cube wrapped in paper.

Personal-Social Behavior: (a) Holds cup to drink from; (b) Inhibits simple acts on command; (c) Repeats performance laughed at.

Normative Summary for Twenty-One Months Level:

Motor Development: (a) Walks attended on the street; (b) Walks backward; (c) Differentiates between stroking and circular scribble.

Language: (a) Joins two words in speech; (b) Names one picture; (c) Repeats things said.

Adaptive Behavior: (a) Places square in form-board; (b) Differentiates between tower and bridge; (c) Folds paper once on demonstration.

Personal-Social Behavior: (a) Bowel control established; (b) Asks for things at table (or for toilet); (c) Pulls persons in order to show something of interest; (d) Tries to turn door knob.

Normative Summary for Thirty Months Level:

Motor Development: (a) Goes up and downstairs alone; (b) Piles seven or eight blocks with coordination; (c) Tries to stand on one foot; (d) Copies vertical or horizontal line.

Language: (a) Points to seven pictures; (b) Names five pictures.

Adaptive Behavior: (a) Attempts to build bridge from model; (b) Adapts to form-board with corrected initial error; (c) Places one completion form; (d) Marks twice for cross.

Personal-Social Behavior: (a) Gives full name; (b) Helps mother to put away things.

(c) *The Bühler Tests for Infants.*—Bühler (30) has published in English tests devised by herself, Hetzer, Wolf and others. These tests were standardized on children in Vienna and have not yet been re-standardized for children in this country. The scale consists of a series of ten tests for each month from two months to eleven months for the first year;

for the second year there are four groups of tests of ten each for each three months of growth.

A mere statement of some of the tests will give the reader some idea of them:

Month II, Test 1	Response to adult's glance
Month III, Test 4	Looking for the source of a sound
Month IV, Test 4	Grasping a touched object
Month VII, Test 8	Taking a toy away from an adult
Month X, Test 7	Imitating ringing the bell
Second Year, Series I, Test 4..	Observing his image in the mirror
Series IV, Test 2	Understanding a command

From this sampling, it can readily be seen that the tests cover all types of reactions. They are somewhat like the Gesell tests in that they try to cover the all-round development of the infant during the first two years of life.

(d) *The Minnesota Pre-School Tests.*—This is a new scale based on an extended study of the test-items available for young children from such standard scales as the Kuhlmann and Stanford Revisions of the Binet and others less well known. A number of new items are included. Two forms of equal difficulty will be available. Separate norms will be provided for the verbal and non-verbal items which are found in about equal proportion in each form. The scale is at present in process of standardization and will shortly be published.

V. MISCELLANEOUS SCALES

(a) *The Porteus Maze Scale.*—This scale differs from all the others so far described inasmuch as it is composed of a single type of performance for all ages, whereas all other scales include a heterogeneous mixture of tests. The type of test selected by Porteus (15) is the maze, and he has constructed a graded series of mazes, which have been standardized for ages 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, and 14. There is one maze for each of these ages. Year 3 is represented by a diamond, the

sides of which are constructed of two parallel lines about ¼ inch apart. The subject is required to trace around the diamond between the parallel lines. The next is a cross constructed on the same plan. Harder tests follow the principle of the maze proper, where there are many blind alleys to penalize the subject for lack of forethought. All the tests are carried out by tracing with a pencil on a printed maze. The subject is allowed a specific number of trials on each maze, and his performance is scored either plus or minus. The mental age is computed from the highest test passed, deducting a year for each lower test failed, and a half year for each lower test passed only on the second trial. The scale, therefore, would be classed as an age scale.

The scale is easy to give, can be given in a relatively short time, and it is decidedly interesting to children. It is also suitable for the testing of non-English-speaking and deaf children, as no language responses are required. The chief limitation, as a general intelligence scale, would seem to be the fact that it is only testing one type of behavior. It is restricted in the range of responses that it calls forth. Nevertheless it is a very valuable supplementary scale.

The author makes a particular plea for the scale on the ground that it is testing "prudent and pre-considered action"—a type of response which, he believes, is lacking in the Binet scale. A measurement of this type of behavior is very desirable, in order to help detect the "conative" type of mental defective who possesses an average amount of "intelligence" as tested by the Binet.

(b) *The Goodenough Drawing Scale.*—This also is a scale limited to one type of behavior, namely the drawing of a man. But Goodenough (26) believes that it gives a fair measure of intelligence; and correlations with the usual validity criteria are fair. The children are merely told to "make a picture of a man. Make the very best picture that you can." No remarks of any kind that might influence the nature of the drawings are allowed. The pictures are then scored by noting the

presence of certain items, e.g., legs present, legs attached to the trunk, nose present, fingers present, etc. No attention is paid to the artistic quality of the drawing, and the author finds that test scores are "relatively unaffected by the type of art instruction given in the primary grades." Mental ages from 3 to 13 can be computed from the norms. The test is recommended for foreign and deaf children, but no method for giving the test to such children seems to have been standardized.

(c) *The Kohs Block Design Scale.*—Kohs (23) has used a set of sixteen color cubes manufactured by the Embossing Company as the basis for a test. A design made up by blocks of different colors is placed before the subject and he is told to reconstruct this design by means of the other blocks at his disposal, within a given time limit. There are seventeen designs in the series, proceeding from very easy to very difficult. Both time and moves are recorded. These are converted into a score, and the score into a mental age.

This scale, like the two preceding scales, the Porteus Maze and the Goodenough Drawing, depends upon one type of reaction for the determination of mental age. For this reason, such scales are generally used as supplementary to the Binet type of scale. They are rarely used as the sole instrument for the determination of mental level.

(d) *The De Sanctis Scale.*—This short scale appeared in 1906 in the Italian literature, and in 1911 in the American (De Sanctis, 11). It appeared, therefore, shortly after the first Binet Scale, but its scope is much more limited. The intention of the author was to use it for the classification of feebleminded children into the three grades of idiots, imbeciles and morons, and to differentiate these three groups from children of normal intelligence. It was not his intention to try to measure the various degrees of intelligence of the normal child.

That these tests differentiate between the feebleminded and normal child, is questioned by Martin (16). It all depends, of course, as to our conception of the upper limits of feeble-

mindedness. Martin made a thorough study and tentative standardization of the De Sanctis tests and found that "morons as a group are successful to a high degree in passing all the tests (of 53 morons, 36 or 68 per cent passed all)." Martin's most important conclusions are that the tests as arranged by De Sanctis are not quite in order of difficulty, that normal children do better than defectives of the same mental age, that the tests as a series to be used by themselves are too verbal, but that they have high value as tests of mentality and should be valuable as supplementary to the Binet or other scales.

The six tests as arranged by De Sanctis may be briefly indicated as follows:

1. Give me a ball (5 balls of different colors).
2. Which is the ball you just gave me? (same 5 balls).
3. Do you see this block of wood? (show cube). Pick out all the blocks like this from the pile on the table (5 cubes, 3 pyramids, 2 parallelopipeds).
4. Do you see this block? (cube). Point out a figure on the form chart that looks like it (show chart). Point out all the squares on the chart as fast as possible (note time, mistakes and omissions—chart has 10 rows—14 in row—squares, triangles and rectangles).
5. (Spread out blocks on table.) How many? Which is largest? Which is farthest away?
6. Do large objects weigh more or less than small objects? Why does a small object sometimes weigh more than a large one? Do distant objects appear larger or smaller or are they really smaller?

De Sanctis' rough method of diagnosis, the accuracy of which Martin questions, is as follows:

(*a*) If second test is failed, mental deficiency is great.
(*b*) If child goes up to fourth test, deficiency is medium.
(*c*) If child succeeds in fifth test, deficiency is slight.
(*d*) If sixth test is passed, there is no defect.

So far the De Sanctis tests have been little used in this country. The more accurately and better standardized Binet and Performance Scales have rendered them somewhat superfluous.

(e) *A Disguised Intelligence Test.*—Snedden (27) has attempted the construction of an intelligence test, which takes the form of an interview with a subject. From this interview an intelligence rating is obtained. The attempt is to make an interview into a reliable instrument for the measurement of intelligence. The Snedden test is based upon word knowledge. In its present form it is not a general interview and could not be used in the usual interview situation. As used by Snedden with high school children the test possesses high reliability and validity, and shows the future possibilities in this direction. There is a need for several such disguised intelligence tests to be used for the measurement of adults when direct overt intelligence tests could not be employed. These would be valuable in interviews between employer and applicant, lawyer and client, physician and patient, social worker and subject, and so on.

(f) *Other Scales.*—In addition to the five scales we have described in this section there are several to be found in the literature of mental testing. Most of these, however, are merely slight variations of the Binet or of other scales and the rest are more in the nature of tentative suggestions.

Into this category would fall Woolley's (15) Scale for Adolescents, Knox's (14) Scale for Testing Immigrants, Mullan's (17) Scale of 100 points, also for testing immigrants, Cornell's (17) Graduated Scale, Squire's (12) Graded Mental Tests, Haberman's (16) Intelligence Examination, and the like. None of them have been widely used nor adequately standardized, and it is, therefore, inadvisable to attempt to describe them here.

VI. SPECIAL SCALES

This section will include a brief mention of scales devised for special types of individuals, who by reason of special defect

manifestly cannot be tested by means of the ordinary scales or in accordance with the usual procedure.

(a) *The Blind.*—The Columbus Point Scale for the Blind by Haines (16) is a revision and adaptation of the Yerkes-Bridges Point Scale. All tests depending on vision have been eliminated and in their place other tests have been substituted. In addition several new tests have been added. Tests 1, 2, 3, 7, 11, 12 and 16 of the Yerkes-Bridges Scale have been dropped. The following have been added:

Naming objects in a basket by touch.
Size—weight illusion.
Comparison of wooden cylinders, 4 and 6 cm. long.
Adaptation board.
Cube suggestion test.
Chooses nicer feeling, (*a*) serge and silk, (*b*) velvet and serge, (*c*) velvet and carpet.
Orientation, r. and l.; N. S. E. W.
Finger tapping—Knox-Pintner lines.
Memory for digits (backwards).

The total maximum points is one hundred. Tentative norms for blind children for each age are given. Up to date 160 blind people have been tested, ranging in age from 6 to adult. On the basis of these results Haines gives median scores for each age.

The Goddard Revision of the Binet-Simon Scale was very early adapted for use with the blind by R. B. Irwin. Much research work was done with this material, which later on was merged into an adaptation of the Stanford Revision for the blind by Hayes. At the present time Hayes' (30) adaptation of the Stanford Revision is the best intelligence scale for the testing of blind subjects. The scale consists of six tests at each year from Year III to Year X inclusive, and then eight tests at Year XII and six tests at Years XIV, XVI and XVIII. For tests requiring vision, other tests have been substituted. The M.A. and I.Q. are calculated in the usual way.

(b) *The Deaf.*—The most useful scale and the one that has been most widely used for testing the mentality of deaf children is the Pintner-Paterson Performance Scale, which has already been described. No special norms for the deaf have been published because the authors believe that the best comparison of the results of a deaf child's performance is with the general standard. Because the tests are equally suited to the deaf as to the hearing, the deaf must be compared with the hearing standards. If, however, we wish to make a diagnosis, we must keep in mind the two or three year difference in mental ability between the deaf and the hearing, and make such allowance as is feasible.

The Drever-Collins Performance Scale and the Porteus Maze Scale, both described above in this chapter, have been used for the measurement of the intelligence of deaf children. The Drever-Collins Scale, like the Pintner-Paterson, was constructed with this purpose specifically in mind.

In the choice of a suitable scale for testing the mental ability of deaf individuals, we must remember that the problem does not consist merely of eliminating tests that are given orally, or of simply changing auditory stimuli to visual stimuli. The change we are forced to make is much more radical than this. It consists of eliminating entirely all language whether spoken or written. Language is for the deaf something extraneous, something artificial, something that corresponds to a school subject in the case of the hearing child, and the difficulty which the deaf encounter in acquiring language makes all language tests useless for the measurement of their intelligence. The problem of the measurement of the deaf is, therefore, much more complex than is the case with the blind, since all oral language tests are suitable for the blind, for they possess the same opportunities as the hearing-seeing child for the acquisition of language.

VII. FOREIGN SCALES

A few words are appropriate here with reference to intelligence scales which have appeared in foreign countries. No claim, however, can be made for any complete survey of this field.

In England early translations and adaptations of the Binet Scale were made by Johnston (10), Winch (14-15) and probably by others. The best adaptation and standardization for British use seems, however, to be the recent revision by Burt (21). This version, the author claims, "adheres more closely to the original procedure of the French authors than any of the published revisions." All the tests, however, have been standardized on English children. They are arranged in order of difficulty, as well as being allocated to appropriate age groups. This results in an unequal number of tests at each age. At Age VI there are 12 tests while at other ages there are only two or three. The mental age of an individual is determined by the total number of tests passed, and this can be immediately read off from a specially devised chart.

Another revision in the English language has been prepared by Phillips (24) for Australia. There are six tests at each age from age 3 to age 15.

In Germany much work with adaptations of the Binet has been reported by Stern (20), Bobertag, Chotzen and others. Stimulated by the work of Binet, Meumann (13) suggested a rather different type of scale having for each age tests of development, tests of intelligence and tests of environment. He worked out a tentative scale, but this idea does not seem to have been followed up. Hetzer and Wolf (28) have published a scale for infants, and Hederschee (20) a tentative scale for deaf children. Hylla (27) gives a general account of testing in Germany.

In the French-speaking countries the Binet Scale has been used largely in the original version of Binet and Simon. Decroly and Buyse (28) give a general survey of all kinds of

intelligence testing. In Russia, Rossolimo began in 1909 to work out a series of tests for different aspects of intelligence which, being scored separately, are then graphically represented on a diagram called a "psychological profile." Kovarsky (27) gives a good account of this method. There are 28 series of tests, each series scored on a ten point scale and the whole series takes from 2½ to 5 hours to give. Vermeylen (no date) has modified the Rossolimo method by materially reducing the number of tests, but keeping the graphic profile.

In Italy some work seems to have been done with adaptations of the Binet. The most radical revision of the Binet is that suggested by Saffiotti (16). This revision is called the Trèves-Saffiotti Revision. They abandon the concept of mental age, because the mental picture of a child may change from age to age. The tests are standardized for each age and at each age three grades of intelligence are chosen, e.g., deboli (backward); medi (average); forti (superior). Tests passed by 60 to 80 per cent at any one age are diagnostic of the backward grade; those passed by 40 to 60 per cent are diagnostic of the average grade; and those passed by 20 to 40 per cent are diagnostic of the superior grade. Practically all of the tests of the Binet Scale are in this manner standardized for these three levels of intelligence for four ages, namely, 6, 7, 8 and ages 9-11 inclusive. Obviously this attempt at re-standardization by these Italian psychologists is a reaction against the comparison of the performance of a child of a given age with the performance of children of different ages. This is probably a valid objection as we have noted in the previous chapter, but the Trèves-Saffiotti Revision seems to be too limited in scope and in the number of grades of intelligence to present a feasible way out of the supposed difficulty.

The Binet Scale is said to have been adapted for use in China, Japan, Sweden, Russia, Turkey, and elsewhere. The writer is not familiar with these adaptations and does not know whether any of them have added anything new to our concept or procedure of intelligence testing. Nor is the writer

aware of scales, other than those already mentioned, for the individual measurement of intelligence, which have appeared in foreign countries.

Conclusion.—We have attempted in this chapter to give a description of the chief scales for individual examination. These are the chief tools that exist at present for the use of the clinical psychologist for the determination of an individual's mental age. All of the scales described are useful. Some of them are better standardized and more accurate than others. Any of them will add to our knowledge of a particular case. The clinical psychologist should be master of all his tools, and not a slave to any one. His skill and expertness will be manifested by his ability to choose the right tool for any particular job, as well as by the manner in which he uses it.

REFERENCES

Arthur, G. (28). "The Re-Standardization of a Point Performance Scale." *J. of Appl. Psych.*, Vol. 12, 278-303.

Binet, A. and Simon, T. *A Method of Measuring the Development of the Intelligence of Young Children.* Trs. by C. H. Town, Lincoln, Ill. No date.

Binet, A. and Simon, T. (16). *The Development of Intelligence in Children.* Trs. by E. S. Kite. Vineland Training School.

Bronner, A. F. et al. (27). *A Manual of Individual Mental Tests and Testing.* Boston.

Bühler, C. (30). *The First Year of Life.* New York.

Burt, C. (21). *Mental and Scholastic Tests.* King, London.

Carroll, H. A. and Hollingworth, L. S. (30). "The Systematic Error of Herring-Binet in Rating Gifted Children." *J. of Ed. Psych.*, Vol. 21, 1-11.

Cornell, C. B. (18). "A Graduated Scale for Determining Mental Age." *J. of Ed. Psych.*, 8, 539-550.

Decroly, O. et Buyse, R. (28). *La Pratique des Tests Mentaux.* Paris, Alcan.

De Sanctis, S. (11). "Mental Development and the Measurement of the Level of Intelligence." *J. of Ed. Psych.*, 2, 498-507.

Drever, J. and Collins, M. (28). *Performance Tests of Intelligence.* Edinburgh.

Gesell, A. (25). *The Mental Growth of the Pre-school Child.* New York.

Gesell, A. (29). *Infancy and Human Growth.* New York.

Goddard, H. H. (11). *The Binet-Simon Measuring Scale for Intelligence.* Revised Edition. Training School, Vineland.

Goodenough, F. L. (26). "A New Approach to the Measurement of the Intelligence of Young Children." *Ped. Sem.*, 33, 185-211.

Goodenough, F. L. (26). *Measurement of Intelligence by Drawings.* Yonkers.

Haberman, V. (16). "The Intelligence Examination and Evaluation." *Psych. Rev.*, 23, 352-379 and 484-500.

Haines, T. H. (16). *Mental Measurements of the Blind.* Psychol. Rev. Monographs. Whole No. 89.

Hayes, S. P. (30). *Terman's Condensed Guide for the Stanford Revision Adapted for the Blind.* Perkins Institution for the Blind, Watertown, Mass.

Herderschee, D. (20). "Teste für Taubstumme Kinder." *Zeitschr. f. Psych.*, 16, 40-61.

Herring, J. P. (22). *Herring Revision of the Binet-Simon Tests.* Yonkers.

Hetzer, H. und Wolf, K. (28). "Babytests." *Zeitschr. f. Psychol.*, 107, 62-104.

Hylla, E. (27). *Testprüfungen der Intelligenz.* Braunschweig, Westermann.

Johnston, K. L. (10). "M. Binet's Method for the Measuring of Intelligence." *Sch. World*, 12, 391 et seq.

Kent, G. H. and Shakow, D. (28). "Graded Series of Form Boards." *Pers. J.*, 7, 115-120.

Knox, H. A. (14). "A Scale Based on the Work at Ellis Island for Estimating Mental Defect." *J. of Am. Med. Asso.*, 62, 742 et seq.

Kohs, S. C. (23). *Intelligence Measurement.* A Psychological and Statistical Study Based upon the Block Design Test. New York.

Kovarsky, V. (27). *La mesure des capacités psychiques.* Paris, Alcan.

Kuhlmann, F. (12). "A Revision of the Binet-Simon System for Measuring the Intelligence of Children." *J. of Psycho-Asthenics.* Mono. Suppl.

Kuhlmann, F. (22). *A Handbook of Mental Tests.* Baltimore.

Martin, A. L. (16). *A Contribution to the Standardization of the De Sanctis Tests.* Training School Bulletin, June.

Memoirs (21). The National Academy of Sciences, Vol. XV, 182-193.

Meumann, E. (13). *Vorlesungen.* Leipzig, Engelmann.

Mullan, E. H. (17). "The Mentality of the Arriving Immigrant." Public Health Bull. No. 90. U. S. P. H. Service.

Phillips, G. E. (24). *The Measurement of General Ability.* An Australian Revision and Extension of the Binet-Simon Scale. Teach. Coll. Press, Sydney, N. S. W.

Pintner, R. and Paterson, D. G. (17). *A Scale of Performance Tests.* New York.

Pintner, R. and Toops, H. A. (17). A Chart for the Rapid Computation of Point Scale Scores. *J. of Delinq.* 2, 209-210.

Porteus, S. D. (15). "Mental Tests for Feebleminded; A New Series." *J. of Psycho-Asthenics,* 19, 200-213.
Saffiotti, F. U. (16). *La Misura Dell' Intelligenza nei Fanciulli.* Societa Romana di Antropologia, Roma.
Snedden, D. S. (27). *A Study in Disguised Intelligence Tests.* T. C. Contribs. to Educ. No. 291, Teach. Coll., Columbia University.
Squire, C. R. (12). "Graded Mental Tests." *J. of Ed. Psych.,* 3, 363-380; 430-443; 493-506.
Squires, P. C. (26). *A Universal Scale of Individual Performance Tests:* Examination Manual. Princeton University Press.
Stern, W. (20). *Die Intelligenz der Kinder und Jugendlichen.* Barth, Leipzig.
Stern, W. and Wiegmann, O. (22). *Methodensammlung zur Intelligenz-pruefung.* Barth, Leipzig.
Stutsman, R. (31). *Mental Measurement of Pre-school Children.* Yonkers.
Terman, L. M. and Childs, H. G. (12). "A Tentative Revision and Extension of the Binet-Simon Measuring Scale of Intelligence." *J. of Ed. Psych.,* 3, 61-74; 133-143; 198-208; 277-289.
Terman, L. M. (16). *The Measurement of Intelligence.* Boston.
Terman, L. M. (17). *The Stanford Revision and Extension of the Binet-Simon Scale for Measuring Intelligence.* Baltimore.
Vermeylen, G. *L'Examen psychographique de l'intelligence.* Anvers, Stockman. No date.
Whipple, G. M. (14-15). *Manual of Mental and Physical Tests.* Two Volumes. Baltimore.
Winch, W. H. (14-15). "Binet's Mental Tests." *Child Study.* Volumes VII and VIII.
Woolley, H. T. (15). "A New Scale of Mental and Physical Measurements for Adolescents." *J. of Ed. Psych.,* 6, 521-550.
Yerkes, R. M., Bridges, J. W. and Hardwick, R. S. (15). *A Point Scale for Measuring Mental Ability.* Baltimore.
Yerkes, R. M. and Foster, J. C. (23). *A Point Scale for Measuring Mental Ability.* (1923 Revision.) Baltimore.

CHAPTER VII

GROUP TESTS

Early Attitude.—A group mental test is a test which can be given to a number of subjects at the same time by a single examiner. The group test is to be contrasted with the individual test, which can only be applied to one individual at a time. It is obvious that, if one examiner can handle a number of subjects at one time, there is a great saving of time and labor. Nevertheless, in spite of the obvious economy of time and effort, it is interesting to note that the group test was slow in arriving and in establishing itself as a legitimate method for the measurement of mental ability. The individual scale was well established long before the advent of the group test.

The early attitude of psychologists towards group tests was decidedly hostile. The possible sources of error that might creep in when examining a group were over-emphasized. It was assumed to be impossible to obtain a valid intelligence rating by group methods and, therefore, there was considerable delay in making the attempt.

The Beginnings.—Although the group method in intelligence testing was late in making its appearance, we can trace back the beginnings of the use of groups of subjects to the laboratory of the experimental psychologist and to the investigations of the school psychologist and educator. There was no attempt in these early experiments to arrive at an intelligence rating. Groups rather than individuals were used in order to collect many data in a short time. The investigation of memory was one of the fields in which the group method was extensively used and for which it was well adapted.

As tests for different mental processes were multiplied, the

group method of testing became popular. The use of this method was also greatly stimulated by the investigations of the educational psychologist in the school room, and also by the growing interest in the relationship between different mental processes studied by the mathematical formulæ for correlation. The large number of tests given by the group method can readily be seen from a study of Whipple's book, "A Manual of Mental and Physical Tests."

The transition from such single group tests to a series of group tests, the results of which should be combined into an intelligence rating, was obvious and natural. Thorndike was among the first to see the advantages of this method and he must certainly be considered the leader in this movement. In the field of business psychology Scott devised group intelligence tests. Pintner (17) used eight group tests and combined the results for each child into a mental age. This was done in order to get a preliminary intelligence rating for the purpose of deciding which children in a particular group should then be more accurately tested by means of individual tests. The tests used included several single group tests which had been used effectively as single tests by Pyle (13) and which had been well standardized by him. Pyle had always used these as separate tests, and Pintner made the next logical step inasmuch as he combined the mental ages obtained by any one child on all the tests into a median mental age which was used as a measure of the child's general intelligence.

The Development in the Army.—A great impetus was given to the construction and use of group tests by the advent of mental testing in the army during the World War in 1917-18. We have noted above the beginnings of the group intelligence test, and the development of this method was merely a question of time. The need and value of such a method were recognized, but it would in all probability have developed slowly, because it would have depended upon the initiative of individual psychologists working more or less in isolation. Their work would have had to overcome a natural amount of inertia

and prejudice in the face of the already well-established methods of individual testing, and they would not have been aided by an overwhelming need for the group method. This overwhelming need presented itself in the army situation. Here the problem was to test thousands of men in a short time. Some sort of a group method was obviously an absolute necessity. Again the army situation enlisted the interest and services of hundreds of psychologists. Here, then, instead of individuals working alone, we find a group of psychologists working in cooperation to construct group tests for rating intelligence.

Under these circumstances it was only natural that the group method should within a year reach a degree of development that it otherwise would have taken many years to attain. All that was known about group intelligence tests was utilized and further development was rapidly made. The nucleus of the tests ultimately constructed was the work of Otis and Terman, who at the time of the introduction of psychological tests into the army were at work upon a group intelligence test.

So successful were these tests in the army that they overcame the general prejudice against the group test method. The work showed the great value of group tests and suggested innumerable fields in which they could be of use. It was only natural, therefore, that after this a number of group tests should make their appearance. It will be impossible here to go into a detailed description of all of these tests. We shall indicate how group and individual tests differ, give a brief description of the more common types of material used in group tests and then describe in a few words the more commonly used tests.

The Difference Between Individual and Group Tests.— The difference between a group test and an individual test lies not only in the number of individuals measured at the same time, but also in the method of testing. In the individual test the examiner is expected to put the subject at his ease, to obtain his undivided attention and to encourage him as much

as necessary. Without departing from the standard procedure of the test, the examiner must nevertheless see that as many incidental factors as possible are favorable to the subject. If the subject refuses to cooperate, gives poor cooperation, or shows nervousness, fatigue or the like, the examiner stops the test and re-tests at some more favorable opportunity.

With the group test, however, such individual adjustment to the test is impossible. Only in a general way can the examiner make sure of the favorable adjustment of the group as a whole, but individuals in the group may not be in the most favorable condition for taking the test. An individual child may be fatigued or frightened or antagonistic, and such conditions may affect his score on the group test. The group test, therefore, is not as pure a measure of intelligence as the individual test. The group test contains in its score not only a measure of the intelligence of the individual, but also a measure of his willingness to cooperate and put forth his best effort. Just because of this, it has proved so useful in school, because willingness to cooperate, willingness to work with the group, is an important aspect of achievement in school work. There is no question here of deciding which is better, the group or the individual test. Each is good for its particular purpose.

Common Types of Material in Group Tests.—Certain types of material are found in many group tests and a description of some of them will be given, because it is impossible to reproduce all the group tests now published in this book. It would not be desirable to do so, even if it were possible.

(a) *Opposites.*—This is one of the oldest and most useful tests. The subject is called upon to respond by writing down or indicating the opposite of a given word, or by deciding whether two words denote opposite or similar ideas.

Underline the word in parenthesis which is the opposite of the first word:

accept. . . . (receive, percept, deny, reject, spend).
constant . . (always, fickle, stationary, seldom, movable).

Underline "opposite," if the two words mean the opposite, and "same" if they mean the same:

furtive	...............sly	same..opposite
any	.................none	same..opposite
asunder	..............apart	same..opposite
deplete	..............exhaust	same..opposite
superfluous	...........essential	same..opposite

If the two words mean the same, write S on the dotted line between them, if they are different write D:

hill ....valley
genuine ....real
useless ....useful
brief ....short

(b) *Analogies.*—This test has proved one of the most valuable. The analogy between two words is given and the subject has to decide as to a similar analogy with reference to another pair.

Underline the best of the four words in parenthesis:

cellar . . . attic:	bottom . . . (well, tub, top, house)
man . . . . arm:	tree. . . . . . (shrub, limb, flower, bark)
imitate. . copy:	invent. . . . (study, Edison, machine, originate)
physics. . motion:	(?). . blood. . (temperature, body, physiology, geography)

Underline the right word:

Coat is to wear as bread is to
eat. . . . starve. . water. . cook
Cat is to tiger as dog is to
bark . . wolf . . . bite . . . snap
Cannon is to rifle as big is to
bullet. . gun. . . . small . . army. . . pistol
Toil is to soil as pay is to
check. . gay. . . . debt. . . money. . work

(c) *Best Reasons.*—This test appears in many forms. It is often called a test of common sense, or comprehension. The

subject indicates in some form or other the best answer to a question.

Check the best reason:

The cause of echoes is

(*a*) the reflection of sound waves.
(*b*) the presence of electricity in the air.
(*c*) the presence of moisture in the air.

Gold is more costly than lead, because

(*a*) it is of finer appearance.
(*b*) it is more scarce.
(*c*) it is used more for jewelry.
(*d*) it is yellow.

The saying, "A carpenter should stick to his bench," means

(*a*) carpenters should not work without benches.
(*b*) carpenters should not be idle.
(*c*) one should work at the thing he can do best.

Why are criminals locked up?

(*a*) to protect society.
(*b*) to get even with them.
(*c*) to make them work.

(d) *Disarranged Sentences.*—This seems to have had its origin in the Binet tests. The words in a sentence are disarranged and the subject has to arrange them properly.

Underline "true" or "false" according to the meaning of the disarranged sentence:

will live bird no forever	true . . false
always sleeplessness clear causes a conscience	true . . false

Cross out the superfluous word in the disarranged sentence:

watch summer the man stole is jail who the in
bushes trees hay roots have and their the ground in.

(e) *Proverbs.*—The subject has to match proverbs of similar meaning, or decide whether they are the same or different in

meaning, or match them with statements which explain the meaning.

Proverbs.

1. The burnt child dreads the fire.
2. Rome was not built in a day.
3. There is no smoke without fire.

etc.

Mark the statements which explain these proverbs:
Time is required to produce anything of value.
Failure follows frequent change of plan.
Unhappy experiences teach us to be careful.
Those in disgrace always want to disgrace others.
There is no result without a cause.

(f) *Number Completion.*—This calls for discovering the rule or method in the arrangement of a series of numbers and indicating this in some way.

Write down the two numbers that should come next:

3 4 6 9 13 18 ___ ___
21 18 16 15 12 10 ___ ___

Fill in the missing numbers:

26 22 ___ 14 10 ___ 2
72 ___ ___ ___ ___ 37 ___ ___ ___ ___ 2

Cross out the number that does not belong in the series:

2 4 8 10 16 32
72 36 18 9 6

(g) *Directions.*—This is one of the earliest types of group test. The subject is asked to do just what he is told:

Cross out the "g" in tiger.
Put a dot below this line —.
If Decoration Day comes in winter write the word "No." If not, write the word "Yes."
Write the letter which follows the letter which comes next after C in the alphabet.

(h) *Sentence Completion.*—Words omitted in a sentence or passage have to be filled in.

Write one word on each blank:

The boy ____ two dollars to the Red Cross.
Those things _____ _____ not fear _____ sometimes _____ harmful.
Poverty cannot ____ down a man ____ is intelligent and ____ hard.

Underline the right word of the three corresponding to the number in the blank:

Once upon a (1) there was a young (2) who was very (3) . He went from (4) to (4) trying to find (5) .

1.	time	place	cat
2.	bird	man	woman
3.	rich	strong	poor
4.	place	there	snow
5.	him	gold	work

(i) *Information.*—The subject is required to show his general information. Usually the items are spread over a wide field. In so far as the material is based on school knowledge, it is an attempt to get at the intelligence of an individual by measuring what has been learned and retained. This may at times give us wrong measures, particularly if we compare children of very different environments.

Underline the correct word:

Euchre is played with dice, rackets, cards, pins.
The Delco System is used in plumbing, filing, ignition, cataloguing.
Coral is found in trees, reefs, molluscs, mines.
John Wesley was famous in literature, science, war, religion.

(j) *Arithmetical Problems.*—Reasoning problems in arithmetic are offered for solution. This test is frankly educational

and does not differ from an achievement test in arithmetical problems.

(k) *Word Knowledge.*—This is tested by asking for the meanings of single words or of words used in sentences.

Underline the word that means the same or nearly the same:

people (1) heavy, (2) story, (3) not now, (4) men and women, (5) open.
kind (1) open, (2) fall, (3) good, (4) not far, (5) new.
pair (1) bag, (2) list, (3) two, (4) yard, (5) party.
action (1) play, (2) deed, (3) mention, (4) opinion, (5) crime.
scrivener (1) searcher, (2) forger, (3) chaplain, (4) clerk, (5) skeptic.
madrigal (1) song, (2) mountebank, (3) lunatic, (4) ribald, (5) sycophant.

"To congregate" means about the same as (1) to dismiss; (2) to assemble; (3) to contradict; (4) to confirm; (5) to applaud.
"To repress" means about the same as (1) to press forward; (2) to press hard; (3) to hurry; (4) to require; (5) to restrain.

Draw a line under the right answer:

Are mysterious things often uncanny?.......Yes No
Is a curriculum intended for horses?........Yes No
Are tentative decisions usually final?........Yes No

(l) *Classification, Generalization.*—There are many tests requiring some kind of generalization, classification or logical selection. This may be tested in several different ways.

In each row draw a line under each of the two words that tell what the thing always has:

Table—books, cloth, dishes, legs, top.
Scissors—cloth, cutting, edge, metal, paper.
Idiocy—crime, foolishness, poverty, stupidity, tuberculosis.

Which one of the five things below is most like these three: horse, pigeon, cricket.

1 stall; 2 saddle; 3 eat; 4 goat; 5 chirp.

Draw a line under the *two* words which tell what the thing always has:

A circle always has: altitude, circumference, latitude, longitude, radius.

Abhorrence always involves: aversion, dislike, fear, rage, timidity.

In each line cross out the word that does not belong there:

Frank, James, John, Sarah, William.
mechanic, doctor, lawyer, preacher, teacher.
give, lend, lose, keep, waste.

Underline the extra word:

football, chess, hockey, basketball, golf.
clarify, explain, argue, illuminate, elucidate.

Draw a line through one word that does not belong with the others:

robin, geranium, poppy, elephant, bluebird.
pulp, knife, bark, root, leaves.
lazy, light, idle, heavy, busy.

(m) *Non-Verbal Material.* Almost all of the verbal types of material are duplicated in some form or other in non-verbal material. Most of this type of material cannot be conveniently reproduced here. Substitution or code tests are very common. All kinds of symbols may be used. The picture completion test calls for the addition of the omitted part. The picture absurdity test requires the subject to indicate what part of the picture is absurd. The maze test requires the correct marking of the shortest passage. There are many tests which require the subject to copy geometrical forms. The cube test requires the subject to calculate the number of cubes from pictures of piles of cubes. The series completion is carried out by symbols instead of numbers. Picture sequence requires the subject to number a series of pictures so that they will form a logical sequence. Similar pictures require the discrimination between similarities and opposites. Analogies are represented by pic-

tures, the subject having to choose from several pictures that one which makes the best fourth picture. Æsthetic judgment requires the marking of the prettiest of three or more pictures or diagrams. Marking a thing shown or a class of things among many other pictures is also used as a test. Marking a given number of things tests ability to count. The dot imitation test requires the subject to draw lines from one dot to another in accordance with the movements of a pointer. The picture arrangement test demands the rearrangement of the parts of a picture. To become familiar with all this varied non-verbal material, the reader must study the various tests themselves. No verbal description is adequate.

DESCRIPTION OF GROUP TESTS

It is obviously impossible in a book of this kind to reproduce all the tests. A short description of the most commonly used and readily available group tests will be given. The description will begin with those useful for the primary grades and will proceed to the more difficult tests suitable for High School and for College.

1. *Pintner-Cunningham Primary Mental Test.*—This is a non-verbal test for kindergarten, first and second grades. The sub-tests consist of: (1) Common observations; (2) Æsthetic judgment; (3) Associated objects; (4) Discrimination of size; (5) Picture parts; (6) Picture completion; (7) Dot drawing. No knowledge of letters and numbers is required. Age and grade norms are available. Figure 15 shows a page of sub-test 2, and figure 16 shows a page of sub-test 6. The reliability of the test for a grade group taking odd and even items is about .9. It correlates with Binet Tests from .6 to .88 (Pintner, 27). It has been compared with many other primary tests by Sangren (29), McGraw and Mangold (29), Dougherty (28), Kuhlmann (28), Heckman (24) and others.

2. *Pintner Non-Language Primary Test.*—As the name implies, this test requires no language for understanding direc-

tions or working the test. It is, therefore, very useful for testing children on their entrance to school, especially with groups

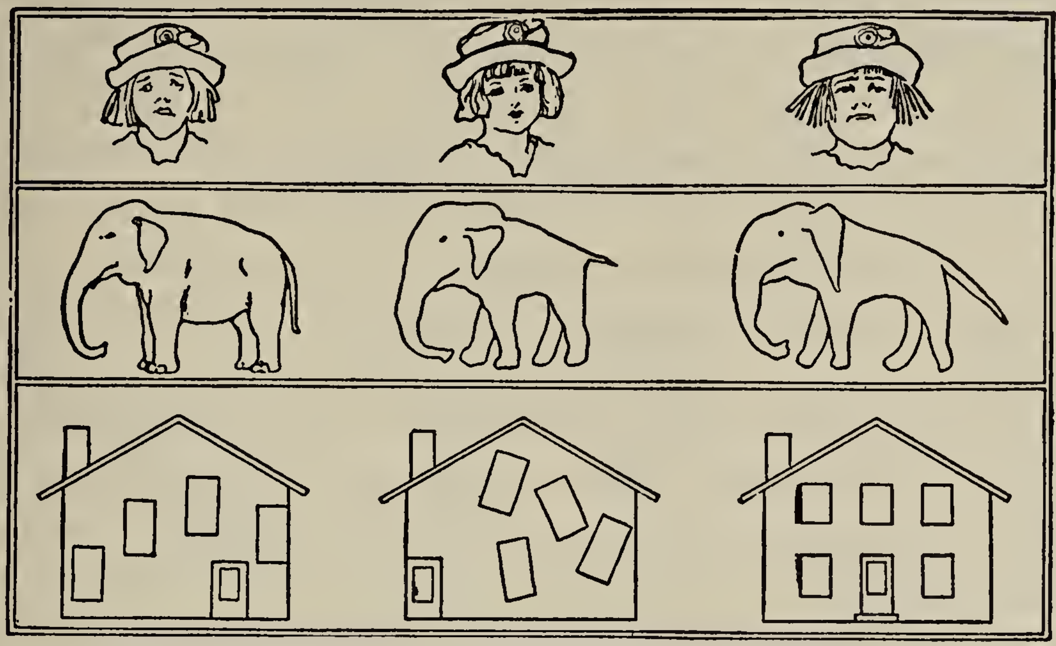


FIG. 15.—A Page of the Pintner-Cunningham Primary Mental Test. Mark the prettiest in each row. (World Book Co.)

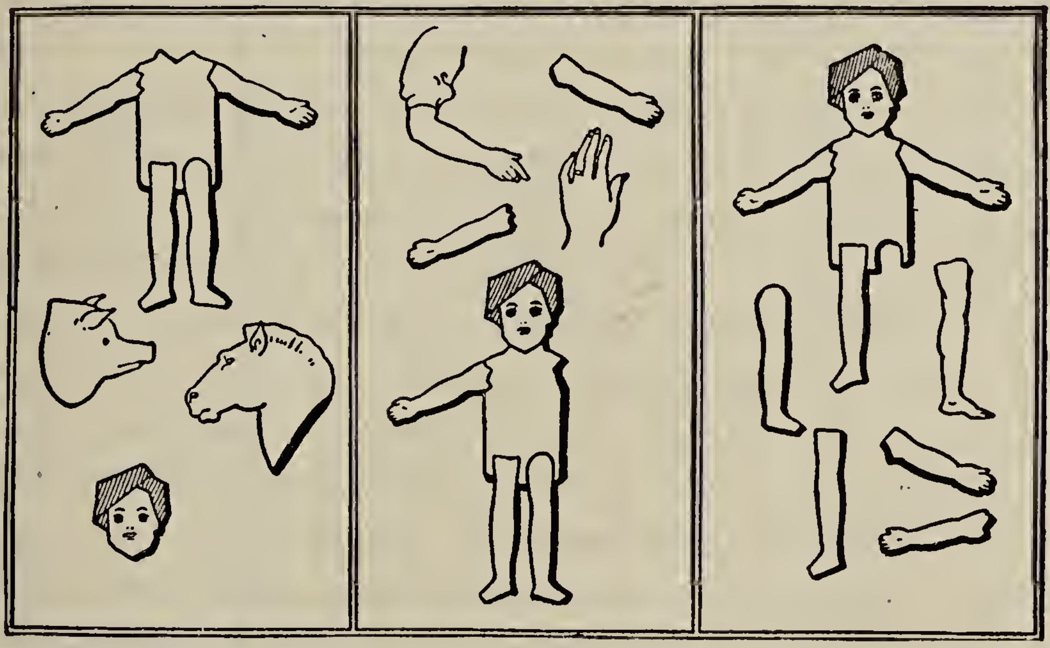


FIG. 16.—A Page of the Pintner-Cunningham Primary Mental Test. Find the missing part. (World Book Co.)

made up of different racial and social backgrounds. It is much more difficult to give than the usual test with verbal directions, but it undoubtedly reduces the errors made by non-verbal tests with verbal directions when given to children having different

amounts of language comprehension. Its sub-tests consist of (1) marking objects held up before the child; (2) completing unfinished geometrical forms; (3) completing unfinished faces; (4) manikin test for position of arms. Figure 17 shows a page from sub-test 1, and figure 18 a page from sub-test 4. The reliability of the test, taking odd and even items, is for Grade I .90 and for Grade II .86. A correlation of .61 with Stanford-Binet for 80 kindergarten children, and a correlation of .51 with the Pintner-Cunningham Test for 154 cases have been reported.

3. *Detroit First Grade Intelligence Test.*—This scale consists of 15 short tests all of which are of the picture type. These tests include following directions, picture completion, drawing geometrical figures, counting, picture-symbol, dot patterns, marking absurd pictures, and others. The tests are easy to give and fairly easy to score. They are suitable for children entering school, and norms for such children are given. The norms are based on over ten thousand children and letter ratings (A, B, C +, C, C —, D, E) are used.

4. *The Pressey Primer Scale.*—This scale consists of four tests, requiring in all four the same type of response, namely, the crossing out of some superfluous member. In this respect the principal type of response demanded of the child is easily learned and the special directions for each test are much reduced. This is undoubtedly an advantage in the administration of the test. Whether this similarity in response from test to test limits the scope of the examination and, thereby, reduces the differentiating power of the test, is a very natural question. This type of examination may favor too much one type of ability to the disadvantage of others in the complex known as general intelligence. These are, however, theoretical considerations that can be best solved by empirical investigation.

The first test requires the child to cross out the extra dot in a series of dots; the second test the crossing out of the unlike picture in a series of three pictures; the third test the

Fig. 17.—A page of the Pintner Non-language Primary Test. Marking objects held up. (T. C. Bureau of Publications.)

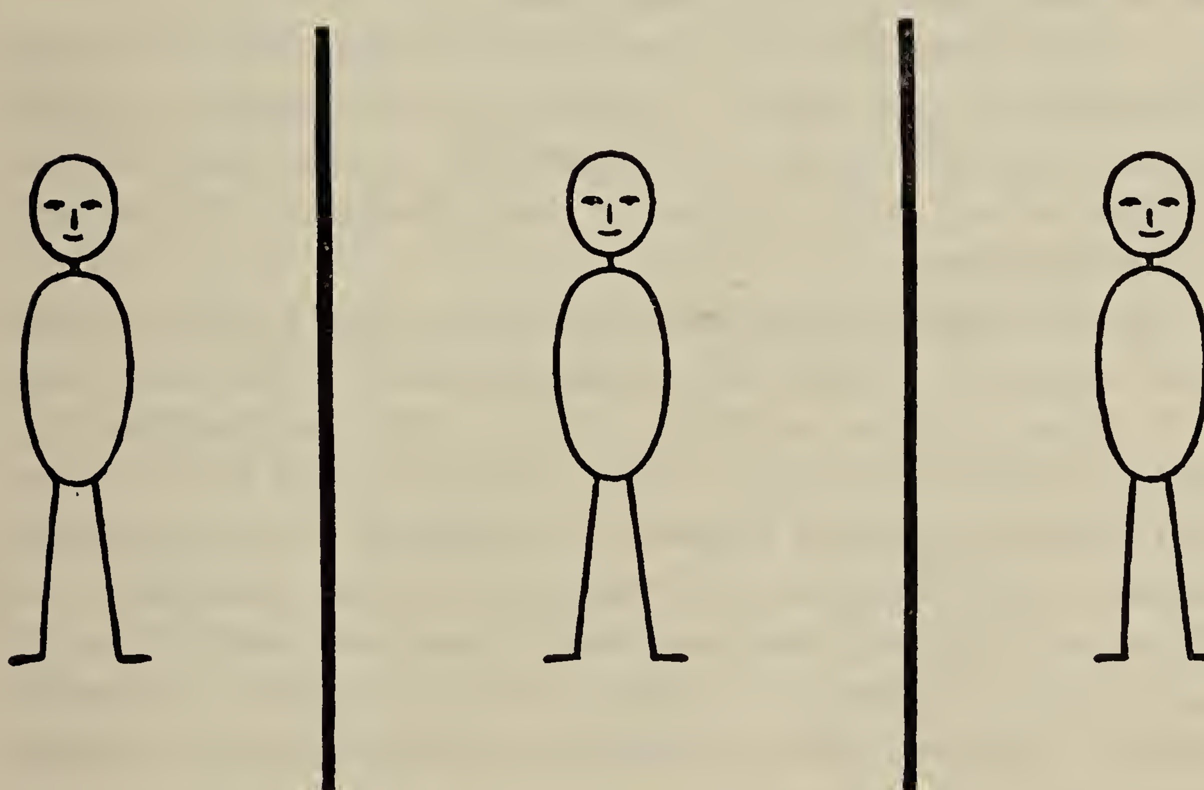


Fig. 18.—A page of the Pintner Non-language Primary Test. Marking in the position of arms of examiner. (T. C. Bureau of Publications.)

crossing out of the extra geometrical form; the fourth test the crossing out of the wrong element in a picture. The directions are given to the children orally. The test is well standardized and excellent norms for first and second graders are available.

5. *Dearborn Intelligence Scale—Series I.*—This is specially adapted to Grades I to III. It is composed of two examinations, A and B. The material is non-verbal in content. The sub-tests are: (1) Directions; (2) Drawing; (3) Counting; (4) General Information; (5) Substitution; (6) Dominoes; (7) Picture Completion; (8) Picture Recognition; (9) Estimation of distances; (10) Number form puzzles. Age norms for ages 5 to 12 are available. The tests are very interesting to children, but some of the items are very difficult to score objectively.

6. *Otis Primary Examination.*—Consists of 8 tests: (1) Following directions; (2) Association; (3) Picture Completion; (4) Maze; (5) Picture Sequence; (6) Similarities; (7) Synonym-antonym; (8) Common sense.

7. *Kingsbury Primary Group Scale.*—This scale is devised for grades 1, 2, 3, and 4. It consists of four tests: (1) Right answers to little stories; (2) Opposites in pictures; (3) Completion series; (4) Form test or Block Building. All the tests are picture tests.

The tests have been given to children in the first four grades and norms for these children are available (Kingsbury, 24).

8. *Rhode Island Intelligence Test.*—This test by Bird (23) is for children from ages 3 to 6. It consists of six sub-tests; (1) Marking common objects; (2) Omissions; (3) Family relations; (4) Common activities; (5) Social groupings; (6) Contrast, similarity and number. There are age medians for ages 3 to 6. Poull (27) finds this test suitable for feeble-minded children. The reliability coefficient for 330 children, ages 3-6, is given as .92.

9. *Detroit Advanced First Grade Intelligence Test.*—This is a non-verbal test with seven sub-tests, as follows: (1) Memory test, marking pictures; (2) Missing parts; (3) Marking objects

described; (4) Detecting similar designs; (5) Marking parts of objects; (6) Counting, marking a given number of objects; (7) Classification, marking things that are alike. Mental age equivalents for ages 5 to 10 are given, as well as letter ratings for ages 5 to 10. The re-test coefficient for 158 pupils is .94.

10. *The Haggerty Delta 1.*—This test is designed for grades 1 to 3. It was devised for and used in the Virginia School Survey. It consists of 12 exercises, six of which are fore-exercises for the purpose of practice, and the other six are the tests proper. Each of the tests is, therefore, preceded by a fore-exercise in the same kind of performance as the test which follows. According to Haggerty, "The fore-exercise is intended to serve two purposes: (1) to afford opportunity for giving preliminary instruction in the method of performing the real test, and (2) to give all pupils some practice in the test in order to equalize the preliminary experiences of the children with the test."

The six tests of the series may be briefly described as follows: (1) Following oral directions; (2) Copying designs; (3) Picture completion; (4) Picture comparison, i.e., deciding whether two pictures are the same or different; (5) Symbol-digit; (6) Word comparison. The first four tests deal with pictures alone. The fifth test introduces numbers, and the sixth, words. These two tests would, therefore, make the whole scale unsuitable for the classification of children entering school for the first time.

11. *The Haggerty Delta 2.*—This test is designed for grades 3 to 9. It is an adaptation of the Army Intelligence Examinations and was devised for and used in the Virginia School Survey. There are six exercises as follows: (1) Discrimination between true and false statements; (2) Arithmetic; (3) Picture completion; (4) Discrimination between words, whether same or opposite; (5) Common sense judgments; (6) General information.

Most of the useful devices used in the Army Tests for ease in scoring, etc., are found in this test. It is better than the

Army Test for school purposes. The norms consist of the average score for each age for ages 8 to 15, and each grade for Grades III to IX (Haggerty, 23).

12. *Dearborn Intelligence Scale.—Series II.*—This is specially adapted to Grades IV to XII. It consists of two examinations, containing the following tests: (1) Picture sequences; (2) Word sequences; (3) Form completion; (4) Opposite completion; (5) Faulty pictures; (6) Disarranged proverbs; (7) Number problems. Norms for ages 6 to 20, and for Grades II to XII are given.

13. *The Pressey Cross-out Tests.*—This group test is useful for Grade III to High School. There are four exercises, each calling for the same type of response, namely, crossing out something. Test 1, cross out the superfluous word in disarranged sentences; Test 2, cross out the superfluous word in lists of words related to each other; Test 3, cross out the superfluous number in a number series; Test 4, a moral judgment test in which the worst thing in the list is to be crossed out. It will be noted that the last test differs radically from the type of test usually included in intelligence examinations. It calls for moral judgments and assumes that a high degree of conformity in these with the conventional standards goes along with high general intelligence. This is probably true, and yet the test seems a little out of place in a general intelligence examination.

There are excellent norms for these tests for ages 10 to 17, and for Grades III to XII.

14. *Pintner Rapid Survey Test.*—This is designed for Grades IV to VIII and has two forms (Pintner, 27). It consists of four sub-tests: (1) Opposites; (2) Analogies; (3) Number Sequence; (4) Classification. Figure 19 shows the Number Sequence Test. All of the items are arranged in multiple-choice form and the subject indicates the answer by writing a number or letter in the margin. The correct numbers or letters form a given sequence (date or word), so that scoring is extremely simple, rapid and accurate. The use of one form of

the test is designed for the rapid survey of schools and school systems where grade or class averages are chiefly desired. The correlation for 55 means of grades between Form A and B is .87. The use of two forms combined is designed for individual purposes. The reliability of two forms is .88. The two forms

NUMBER SEQUENCE TEST

Put letter here

Look at the sample that follows:

2 4 6 8 10 12—

What number should come next? (A) 17 (B) 13 (C) 14 (D) 15 (E) 9

The right answer is 14, so put the letter that goes with it, C, in the margin. → C ....

Read this sample:

1 8 2 8 3 8— (E) 8 (F) 6 (G) 2 (H) 4 (J) 9

The right answer is 4, so you must put H in the margin. → H ....

Do the rest in the same way. Remember to put the letter in the margin.

1	2	3	4	5	6—	(N) 5	(O) 2	(P) 7	(Q) 3	(R) 8	....
40	35	30	25	20	15—	(S) 10	(T) 5	(U) 14	(V) 15	(W) 20	....
8	7	6	5	4	3—	(W) 1	(X) 3	(O) 9	(Y) 2	(Z) 4	....
3	3	5	5	7	7—	(A) 8	(B) 11	(C) 9	(D) 6	(E) 10	....
5	9	13	17	21	25—	(F) 28	(G) 30	(H) 29	(J) 21	(K) 26	....
9	9	7	7	5	5—	(K) 4	(L) 10	(M) 6	(N) 2	(O) 3	....
4	3	5	4	6	5—	(K) 3	(L) 7	(M) 6	(N) 2	(O) 3	....
19	16	14	11	9	6—	(K) 3	(L) 5	(M) 2	(N) 7	(O) 4	....
103	95	87	79	71	63—	(D) 64	(E) 54	(F) 47	(G) 55	(H) 51	....
81	27	9	3	1	1/3—	(X) 2/3	(Y) 1/9	(Z) 0	(A) 1/27	(B) 1/81	....

FIG. 19.—Test 3 of the Pintner Rapid Survey Intelligence Test. (T. C. Bureau of Publications.)

combined correlate .83 with the Terman Group Test for 540 children in Grades IV to VIII, and .78 with the National Intelligence Test for 428 children in Grades III to V.

15. *The National Intelligence Tests.*—These tests were prepared under the auspices of the National Research Council by Haggerty, Terman, Thorndike, Whipple and Yerkes. There are several different forms available. Two booklets are to be used for each examination. Each test booklet contains five exercises. Scale A contains: (1) Arithmetic; (2) Sentence

completion; (3) Checking attributes possessed by a given word; (4) Discrimination of similarity and difference as applied to words; (5) Symbol-digit tests. Scale B contains: (1)

In each row draw a line under each of the two words that tell what the thing always has.

SAMPLES	man	(body cane head shoes teeth)
	dog	(blanket chain collar legs nose)
	house	(cellar paint room servants walls)

Begin here	1	hen	(chickens corn feathers neck roost)
	2	tiger	(bones cage cubs fur jungle)
	3	squirrel	(acorn fur nest peanut tail)
	4	kitten	(ball claws eating eyes mouse)
	5	bicycle	(basket bell brake frame wheels)
	6	stone	(field hardness hurt throwing weight)
	7	lion	(cage head keeper mane prey)
	8	face	(cheeks eyebrow glasses mustache mouth)
	9	forest	(cones flowers grass soil trees)
	10	paper	(edges envelope printing surface watermark)
	11	Borneo	(airplanes land rivers saloons universities)
	12	gully	(flowers sand sides steepness stream)
	13	piano	(keys music pedals scarf stool)
	14	satisfaction	(conquering contentment money pleasure trouble)
	15	illness	(ailment discomfort doctor nurse recovery)
	16	mob	(confusion excitement hanging negro torches)
	17	fire	(ashes danger flame heat wood)
	18	sea	(coast reefs salt shoals submarines)
	19	alley	(cans fence narrowness passage pavement)
	20	crime	(death lawlessness punishment theft wrong)
	21	pilot	(cap chart knowledge license raincoat)
	22	measles	(discomfort doctor nurse rash recovery)
	23	nun	(beauty convent teacher vow woman)
	24	citizen	(city country male privileges vote)

FIG. 20.—A page from The National Intelligence Test. Scale A. (World Book Co.)

Arithmetic; (2) General Information; (3) Logical Judgment; (4) Analogies; (5) Discrimination of similarity and difference as applied to numbers, names and forms. Figure 20 shows the Common Attributes Test.

The novel feature of this test is the fore-exercise that pre-

cedes each exercise proper. This fore-exercise is a sample of the performance to be carried out by the examinee and gives him an opportunity of some practice before beginning the test proper. In most cases the length of the fore-exercise is a little less than half of the test proper. In no other test has this theory of the necessity for preliminary practice been carried to such an extent. It conflicts somewhat with the theoretical definition of general intelligence as being a measure of ability to adapt. Quickness and readiness of adaptation become more difficult to measure the greater the amount of previous practice. This is theoretical merely and it may be that tests with an appreciable degree of fore-exercise will nevertheless prove of higher value than those having little or no fore-exercise. In this case the above interpretation of the definition of general intelligence will have to be revised.

There are age and grade norms available, based on many cases. Kelley (27) estimates the reliability for a single grade range for Scale A to be about .7 and for Scale B about .75 and for A + B to be about .85. The correlation between Scales A and B for 1073 cases, Grades III to VIII, is .93.

16. *The McCall Multi-Mental Scale.*—This test devised by McCall (26) is very different from the usual intelligence test. Figure 21 shows a reduced picture of the test, which consists of one sheet only and has no sub-tests. The subject has to cross out one word in each of the 100 groups of words, after having the procedure carefully explained by means of the five samples at the beginning. What is to be crossed out differs from one group of words to the next. A different mental set is required for each group of words and this ability to shift rapidly from one mental set to another is considered by the author as a good index of intelligence. It correlates .93 with other criteria of intelligence.

17. *The Pintner Non-Language Mental Test.*—This is a test useful in Grades IV to VIII. Each test is demonstrated on the blackboard and no language is necessary for the understanding of the directions or for doing the exercises of the test.

MULTI-MENTAL SCALE

Elementary School
Form 1

By William A. McCall and His Students, Teachers College, Columbia University

Name .... Grade .... School .... City .... Age .... Boy or Girl ....

A	B	C	D	E
chair	gate	sweet	mama	dog
dog	good	ripe	sister	leaps
table	big	red	papa	stone
bed	bad	crow	brother	runs
stove	little	apple	grandma	barks

1	2	3	4	5	6	7	8	9	10
fly	cup	horse	lesson	grass	robin	high	irrigate	black	word
burn	fork	calf	problem	coal	geranium	low	land	hot	paragraph
gasoline	saucer	colt	teacher	carbon	elephant	cat	soil	white	sentence
coal	bowl	hen	learn	tar	poppy	fever	cultivate	star	style
wood	knife	cow	solve	soot	bluebird	dangerous	navigate	cold	composition

11	12	13	14	15	16	17	18	19	20
baby	another	cruel	vote	foreign	gold	moon	stone	needle	quickly
slow	first	cheerful	decide	brave	steel	seed	wood	scissors	buy
donkey	either	courteous	citizens	rose	ore	tree	iron	paper	manufacture
gate	last	lad	factory	coin	spring	root	horse	cloth	shoes
sleepy	neither	generous	juries	fragrant	iron	sapling	clay	thread	sell

21	22	23	24	25	26	27	28	29	30
lesson	grapes	cloud	mine	wind	stockings	violin	murmurs	fence	Washington
holiday	plums	excited	water	blows	shirt	flute	smiles	rock	Lee
study	peaches	water	cap	lightning	mend	guitar	brooklet	lime	Grant
learn	raisins	running	lake	thunder	head	mandolin	babbles	mortar	Lincoln
recite	prunes	crowd	iron	flashes	coat	banjo	sparkles	wall	Edison

31	32	33	34	35	36	37	38	39	40
half	selfish	in	desk	flock	bird	drums	cat	eye	speed
seventh	sincere	out	teacher	cattle	swim	lesson	bird	nose	interest
nickel	man	on	pew	barn	crawl	sharpened	fins	chin	legislature
quarter	kind	up	school	geese	fish	beat	fly	hand	money
dime	generous	under	church	herd	snakes	studied	fish	ear	laws

41	42	43	44	45	46	47	48	49	50
chair	tire	come	sheep	write	pulp	baby	mother	knife	barometer
letter	engine	cube	rose	read	knife	boy	student	sword	government
books	horn	edge	busy	book	bark	man	book	dagger	establish
poems	mast	face	beautiful	house	root	works	teacher	shield	investigate
write	wheel	square	bee	letter	leaves	cries	daughter	saber	criticize

51	52	53	54	55	56	57	58	59	60
hat	fruit	rain	lines	arrow	brick	owns	elect	boy	gasoline
kind	made	snow	draw	pointed	frame	sells	president	girl	wagon
cheerful	money	storm	coincides	sticky	concrete	house	king	kitten	streetcar
quick	bread	wind	points	swift	silver	sits	appoint	run	auto
boy	gave	hail	angles	straight	houses	buys	secretary	apple	electricity

61	62	63	64	65	66	67	68	69	70
door	floor	perfume	nation	man	lazy	cat	picture	corn	automobile
lamp	woman	fragrant	senate	horse	light	bird	wall	wheat	aeroplane
knob	smoke	tropical	county	rooster	idle	stone	horse	muffins	furnace
table	engine	rare	village	woman	heavy	parrot	carriage	biscuit	radio
boy	men	flowers	state	hen	busy	feed	bug	sausage	telephone

71	72	73	74	75	76	77	78	79	80
leaf	circle	dance	boy	air	runs	knife	army	tell	perseverance
root	cube	bear	barks	earth	rabbit	saucer	camp	language	faithfulness
ground	area	beaver	black	water	dog	spoon	scouts	story	honesty
stem	acre	hunt	dog	fish	jumps	fork	row	well	conflagration
flower	triangle	fight	plays	bird	barks	cup	hike	speak	courage

81	82	83	84	85	86	87	88	89	90
sour	gun	like	paper	large	cat	wrong	swim	churn	Alaska
sweet	hammer	nation	cover	air	monkey	give	spiders	sew	Mexico
bitter	sharp	position	door	white	baboon	right	fish	write	United States
acid	shoot	rule	book	ball	tiger	night	lamb	butter	France
tart	knife	accept	leather	hard	elephant	sleep	spin	dress	Canada

91	92	93	94	95	96	97	98	99	100
eat	books	wool	sweet	chair	smooth	girl	ducks	baby	investigate
sing	power	cloth	lemons	room	road	walk	paddle	puppy	publish
book	knowledge	shoes	cake	hall	great	does	geese	kitten	editors
apple	paper	meat	sour	building	rough	sleep	fish	pig	write
read	food	leather	salty	door	table	play	swim	calf	printing

If you finish before time is called, go back and improve your work.

Published by Bureau of Publications, Teachers College
Columbia University, New York City

Total	
Minus	
Score	

FIG. 21.—The McCall Multi-mental Scale, greatly reduced in size. (T. C. Bureau of Publications.)

The test blank consists of the following six exercises: (1) Movement Imitation, i.e., reproducing the movements of a pointer after it has been moved from dot to dot in different ways on the blackboard. This is essentially the Knox Cube Test arranged for group purposes. (2) Easy Learning, i.e., a very simple digit-symbol test containing three elements. (3) Hard Learning, i.e., a more difficult digit-symbol test containing nine elements. (4) Drawing Completion, i.e., drawing in the missing parts of pictures. (5) Reversed Drawings, i.e., reproducing geometrical forms as they would be when turned upside-down. (6) Picture Reconstruction, i.e., indicating by numbers the positions of the parts of pictures so as to make a complete picture. Figure 22 shows one of the pages of this test where the child has to supply the missing parts of the pictures.

The test takes about 30 minutes to give, and has proved very serviceable, requiring, as it does, simply a blackboard and one demonstration picture. It has been well standardized on ordinary school children. The correlations with various verbal tests for various groups range from .25 to .72 (Pintner, 24). Its validity coefficient with a composite criterion of intelligence is .78. This test has been used frequently for the comparison of native and foreign groups. It has been extensively used for the testing of deaf children and very good norms for such cases are available.

18. *The Otis Group Intelligence Scale.*—This group test is suitable for Grades V to XII. It is also difficult enough for College Students. It consists of ten parts, as follows: (1) Following printed directions; (2) Opposites; (3) Disarranged Sentences; (4) Matching Proverbs; (5) Arithmetic; (6) Geometric Figures; (7) Analogies; (8) Similarities; (9) Narrative completion; (10) Memory. The reliability coefficient for Grades IV to VIII is given as .97.

19. *Otis Self-Administering Test of Mental Ability.*—There are two levels of this test, the Intermediate Examination for Grades IV to IX and the Advanced Examination for High

Schools and Colleges. The self-administering feature of the test refers to the fact that the subject can read over the directions on the first page of the test, and these directions give

Fig. 22.—The Picture Completion Test of the Pintner Non-language Mental Test.

samples of all the different kinds of items which appear in the test proper. The test is not divided into sub-tests, but different types of items appear mixed up throughout the test, beginning with easy items and proceeding to more difficult

ones. There are 75 items in each examination. The tests are very easy to give and to score. The advanced examination is particularly good for the testing of educated adults, not in college, who sometimes resent the oral directions and time limits for each sub-test, which are customary with most intelligence tests. The reliability coefficient for the Intermediate Examination for Grades IV to IX is .95 and for the Advanced Examination for Grades VII to XII is .92.

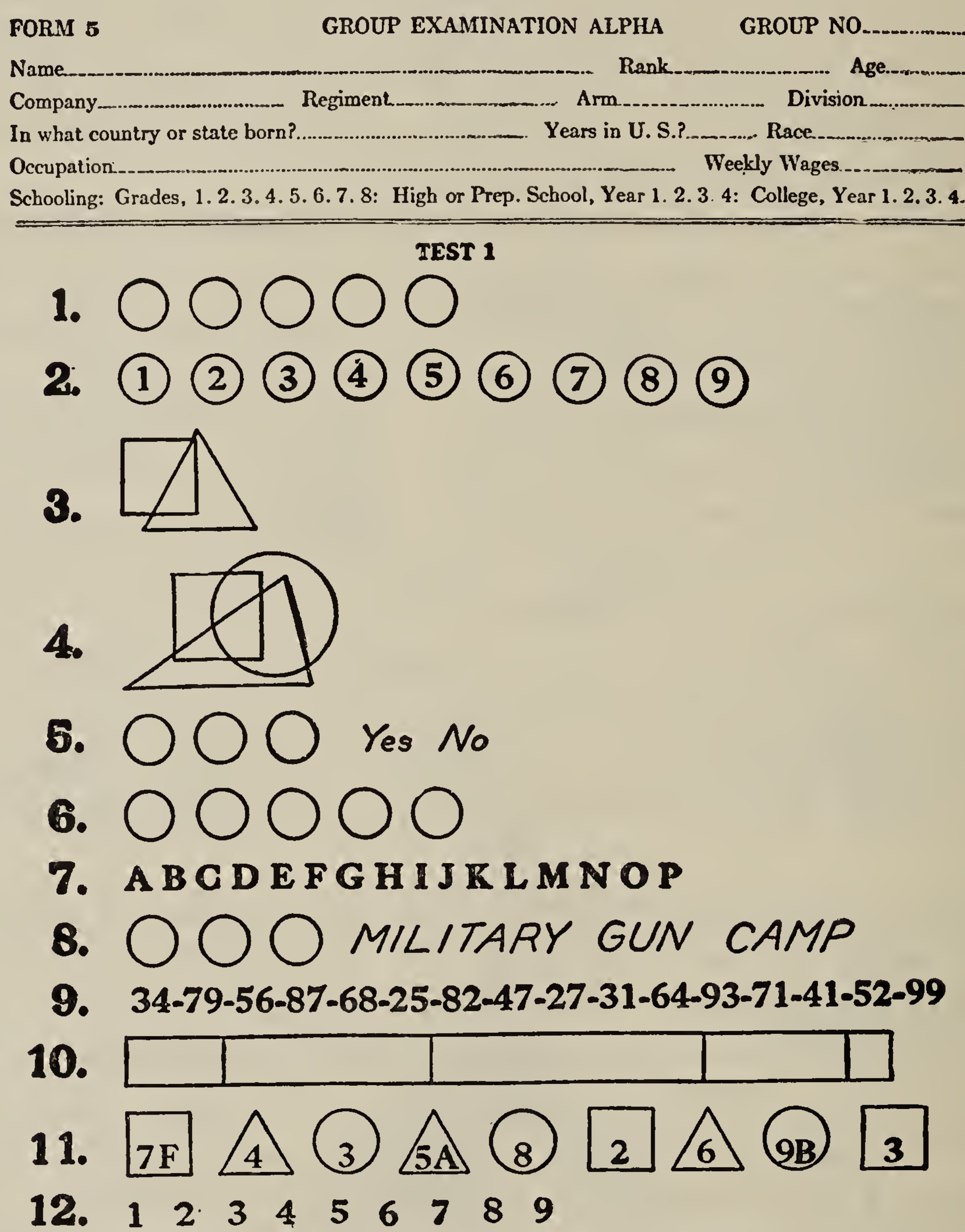
20. *The Army Alpha.*—This test is suitable for adults and was used in the army. It has been effectively used in High School and to some extent in the Elementary School. It proved particularly valuable in the army, but it is not, therefore, the best for use in colleges or schools. Much of the content is so constructed as to appeal to soldiers. There are eight tests as follows: (1) Following directions; (2) Arithmetical problems; (3) Practical judgment; (4) Synonym-antonym; (5) Disarranged sentences; (6) Number series; (7) Analogies; (8) Information. The army tests are reproduced in the "Memoirs" of the National Academy of Sciences (21) and in a book by Yoakum and Yerkes (20). Figure 23 shows the Directions Test of the Army Alpha and Figure 24 shows the Analogies Test.

There are five different forms of the Army Alpha, all roughly of the same degree of difficulty. It is, therefore, extremely useful in testing groups where there is danger of coaching. As this test was given to more than one million unselected recruits, there are very reliable norms for adults. No other group test has so far been given to such a random sampling of American adults, and the standards for this test represent, therefore, the most reliable measures of the intelligence of American adults. In addition, equivalent scores for the Alpha and Stanford-Binet have been computed, so that we may obtain an approximate mental age from the scores on the Army Alpha.

21. *Revision of Army Alpha.*—The Psychological Corporation has prepared a revision of the Army Alpha. This is more suitable for civilians than the original form. Bregman (26)

has prepared percentile equivalents in order to interpret the scores in the light of the total distribution of the army results.

22. *Detroit Alpha Intelligence Test.*—This is a group test


FORM 5 GROUP EXAMINATION ALPHA GROUP NO.........

Name........ Rank........ Age........

Company........ Regiment........ Arm........ Division........

In what country or state born?........ Years in U. S.?........ Race........

Occupation........ Weekly Wages........

Schooling: Grades, 1. 2. 3. 4. 5. 6. 7. 8: High or Prep. School, Year 1. 2. 3. 4: College, Year 1. 2. 3. 4.

TEST 1

1.

2. 1 2 3 4 5 6 7 8 9

3.

4.

5. Yes No

6.

7. A B C D E F G H I J K L M N O P

8. MILITARY GUN CAMP

9. 34-79-56-87-68-25-82-47-27-31-64-93-71-41-52-99

10.

11. 7F 4 3 5A 8 2 6 9B 3

12. 1 2 3 4 5 6 7 8 9

FIG. 23.—Directions Test.—The Army Alpha Intelligence Test.

suitable for Grades V to IX. It is made up of eight sub-tests: (1) Information; (2) Opposites; (3) Classification; (4) Block Design; (5) Generalization; (6) Analogies; (7) Number Sequence; (8) Mixed-up Sentences. There are age norms from age 8 to 18, and letter ratings for ages 9 to 16. The correla-

tion for 51 sixth-grade pupils for two forms is .91; for 273 sixth-grade pupils .85.

TEST 7

SAMPLES
sky—blue : : grass— **table** **green** **warm** **big**
fish—swims : : man— **paper** **time** **walks** **girl**
day—night : : white— **red** **black** **clear** **pure**

In each of the lines below, the first two words are related to each other in some way. What you are to do in each line is to see what the relation is between the first two words, and underline the word in heavy type that is related in the same way to the third word. Begin with No. 1 and mark as many sets as you can before time is called.

1 shoe—foot : : hat— **kitten** **head** **knife** **penny** 1
2 pup—dog : : lamb— **red** **door** **sheep** **book** 2
3 spring—summer : : autumn— **winter** **warm** **harvest** **rise** 3
4 devil—angel : : bad— **mean** **disobedient** **defamed** **good** 4
5 finger—hand : : toe— **body** **foot** **skin** **nail** 5

6 legs—frog : : wings— **eat** **swim** **bird** **nest** 6
7 chew—teeth : : smell— **sweet** **stink** **odor** **nose** 7
8 lion—roar : : dog— **drive** **pony** **bark** **harness** 8
9 cat—tiger : : dog— **wolf** **bark** **bite** **snap** 9
10 good—bad : : long— **tall** **big** **snake** **short** 10

11 giant—large : : dwarf— **jungle** **small** **beard** **ugly** 11
12 winter—season : : January— **February** **day** **month** **Christmas** 12
13 skating—winter : : swimming— **diving** **floating** **hole** **summer** 13
14 blonde—light : : brunette— **dark** **hair** **brilliant** **blonde** 14
15 love—friend : : hate— **malice** **saint** **enemy** **dislike** 15
16 egg—bird : : seed— **grow** **plant** **crack** **germinate** 16
17 dig—trench : : build— **run** **house** **spade** **bullet** 17
18 agree—quarrel : : friend— **comrade** **need** **mother** **enemy** 18
19 palace—king : : hut— **peasant** **cottage** **farm** **city** 19
20 cloud-burst—shower : : cyclone— **bath** **breeze** **destroy** **West** 20

21 Washington—Adams : : first— **president** **second** **last** **Bryan** 21
22 parents—command : : children— **men** **shall** **women** **obey** 22
23 diamond—rare : : iron— **common** **silver** **ore** **steel** 23
24 yes—affirmative : : no— **think** **knowledge** **yes** **negative** 24
25 hour—day : : day— **night** **week** **hour** **noon** 25

26 eye—head : : window— **key** **floor** **room** **door** 26
27 clothes—man : : hair— **horse** **comb** **beard** **hat** 27
28 draw—picture : : make— **destroy** **table** **break** **hard** 28
29 automobile—wagon : : motorcycle— **ride** **speed** **bicycle** **car** 29
30 granary—wheat : : library— **read** **books** **paper** **chairs** 30

31 Caucasian—English : : Mongolian— **Chinese** **Indian** **negro** **yellow** 31
32 Indiana—United States : : part— **hair** **China** **Ohio** **whole** 32
33 esteem—despise : : friends— **Quakers** **enemies** **lovers** **men** 33
34 abide—stay : : depart— **come** **hence** **leave** **late** 34
35 abundant—scarce : : cheap— **buy** **costly** **bargain** **nasty** 35

36 whale—large : : thunder— **loud** **rain** **lightning** **kill** 36
37 reward—hero : : punish— **God** **everlasting** **pain** **traitor** 37
38 music—soothing : : noise— **hear** **distracting** **sound** **report** 38
39 book—writer : : statue— **sculptor** **liberty** **picture** **state** 39
40 wound—pain : : health— **sickness** **disease** **exhilaration** **doctor** 40

FIG. 24.—Analogies Test. The Army Alpha Intelligence Test.

23. *Army Beta.*—This consists of seven tests, all of which are explained in gesture by the examiner and demonstrated on a blackboard chart by the demonstrator. The procedure is acted out before the group.

The tests are: (1) Maze Drawing; (2) Cube Analysis; i.e., counting the number of cubes in drawings of various arrangements of cubes; (3) X-O Series or completing series of crosses and circles arranged in various rhythmic sequences; (4) Digit-Symbol; (5) Number Checking; (6) Drawing Completion; (7) Geometrical Construction, i.e., drawing the divisions in a square to correspond to the separate scattered pieces.

The Beta test was given to thousands of non-English-speaking recruits in the army and proved a valuable means of measuring their intelligence, an end which could never have been attained by means of the usual language group test.

24. *Terman Group Test of Mental Ability.*—This consists of ten tests as follows: (1) Information; (2) Best answer; (3) Word meaning; (4) Logical selection; (5) Arithmetical problems; (6) Sentence meaning; (7) Analogies; (8) Mixed sentences; (9) Classification; (10) Number Series. Suitable for Grades VII to XII. Age and grade norms are available. The reliability for 132 cases in Grade IX is .89. This test is one of the most frequently used tests for high school purposes. It is a little too easy for superior pupils in the last year of high school. Figure 25 shows the Analogies Test.

25. *Miller Mental Ability Test.*—Consists of three tests: (1) Disarranged sentences combined with directions; (2) Controlled association; (3) Analogies. This is a short test and seems to work very well. It is suitable for Grades VII to XII. The reliability coefficient for re-testing 109 pupils in Grade X is .91.

26. *Detroit Advanced Intelligence Test.*—This consists of eight sub-tests: (1) Information; (2) Opposites; (3) Classification; (4) Number Sequence; (5) Block Designs; (6) Spelling; (7) Analogies; (8) Mixed-up Sentences. Norms are given for ages 9 to 23 and letter ratings for ages 11 to 16. The test is designed for high school and college.

27. *Thurstone Psychological Examination.*—This is described as suitable for High School Seniors and College Freshmen, but it can be used effectively for all years in high school

ANALOGIES

Samples
- Ear is to hear as eye is to
  table see hand play
- Hat is to head as shoe is to
  arm coat foot leg

Do them all like samples.

---

1 Coat is to wear as bread is to
eat starve water cook 1

2 Week is to month as month is to
year hour minute century 2

3 Monday is to Tuesday as Friday is to
week Thursday day Saturday 3

4 Tell is to told as speak is to
sing spoke speaking sang 4

5 Lion is to animal as rose is to
smell leaf plant thorn........... 5

6 Cat is to tiger as dog is to
wolf bark bite snap............. 6

7 Success is to joy as failure is to
sadness luck fail work 7

8 Liberty is to freedom as bondage is to
negro slavery free suffer 8

9 Cry is to laugh as sadness is to
death joy coffin doctor 9

10 Tiger is to hair as trout is to
water fish scales swims........... 10

11 1 is to 3 as 9 is to
18 27 36 45........................ 11

12 Lead is to heavy as cork is to
bottle weight light float 12

13 Poison is to death as food is to
eat bird life bad 13

14 4 is to 16 as 5 is to
7 45 35 25......................... 14

15 Food is to hunger as water is to
drink clear thirst pure 15

16 b is to d as second is to
third later fourth last 16

17 City is to mayor as army is to
navy soldier general private 17

18 Here is to there as this is to
these those that then 18

19 Subject is to predicate as noun is to
pronoun adverb verb adjective..... 19

20 Corrupt is to depraved as sacred is to
Bible hallowed prayer Sunday 20

Right........

Fig. 25.—The Analogies Test from the Terman Group Test of Mental Ability. (World Book Co.)

or college. It consists of a great number of problems involving analogies, number completion, logical reasoning, mental arithmetic, general information, sentence completion, proverb matching and the like. Unlike the usual group test, the items are not grouped together in various tests, but are thoroughly mixed up in a spiral arrangement, the same type of problem occurring again and again, beginning with the easiest examples and gradually becoming harder and harder. The subject is given a specific time for the whole examination.

28. *Thorndike Intelligence Examination.*—This is a team of five tests devised by Thorndike and represents the most extensive and thorough intelligence examination that has so far been devised. It requires 2 hours and 50 minutes of working time, and is also relatively difficult to score. Many of the tests require thorough knowledge of the material and sound judgment in order to achieve uniform and accurate scoring.

There are four booklets. The first is the practice form which presents samples of the various kinds of items in the test proper, and fifteen minutes' practice with this is allowed. Then follows Part I, time limit 45 minutes, made up of nine subtests, consisting of directions, arithmetic problems, information, opposites, word meanings, and the like. Part II then follows with a time limit of 50 minutes. It consists of sentence completion, algebraic problems and information. Part III then follows with a time limit of 60 minutes. This consists of answering questions based on difficult reading passages. Frequency distributions for college groups are given, and the significance for future college work for certain scores is indicated. Wood (23) gives a thorough discussion of the uses of this test in higher education. New forms of the test are being constantly made. It is used for the examination of candidates at Columbia College.

29. *Ohio State University Psychological Test.*—This test is used by the Ohio College Association in many colleges in Ohio. New forms are prepared every year. One form of the test consisted of five sub-tests: (1) Mathematical Reasoning; (2)

Same-Opposites; (3) Analogies; (4) Number Sequence; (5) Reading Passages. Very adequate norms for college freshmen are available.

In the most recent form of this test (Form 17), Toops has introduced a rapid scoring device. As all of the responses to the test questions are of the multiple-choice type, the student marks his answer in one of the five numbered little squares. These responses are recorded on an envelope, separate from the test booklet proper, which slides back and forth on the cardboard back cover page of the test. After the test is administered, the perforated ends of the envelope are torn off, leaving a standard size file folder with the answers on the uppermost side of the large sheet when the folder is opened out. The sheet containing the subject's responses is sent to the printer, who, by means of a suitable electrotype plate, so prints over these answer cards that red appears on the four wrong squares of each question and white is left on the correct square. The subject's correct responses, therefore, appear as pencil crosses in little white squares. The number of correct responses can, therefore, be very quickly counted. After scoring, the test responses are folded inside the folder for permanent preservation together with the other personnel papers filed therein, while the front and back covers contain answers to personnel questions and a scholastic record blank respectively. Inasmuch as the examinee does no writing on the test blank itself, the test may be used over and over again by simply replacing the envelope.

Tests Covering Wide Ranges of Mental Ability.—All the tests so far described have been sharply limited in the range of mentality tested. There are several tests which attempt to cover much wider ranges. This is very desirable, but it is difficult to accomplish.

30. *The I. E. R. Intelligence Scale CAVD.*—This is a scale constructed by Thorndike (27) of the Institute of Educational Research, Teachers College, Columbia University. The letters CAVD refer to the four kinds of content used in the tests,

namely, Completion, Arithmetic, Vocabulary, and Directions. There are seventeen groups of tests each measuring a level of intellect, from level A to level Q. Each level is represented by forty items or tasks, ten each of the four kinds of material mentioned above, i.e., ten completion items, ten arithmetic items, and so forth. The lowest level A is suitable for an average mentality of three and then the levels proceed up to level Q which is difficult enough to measure the intelligence of the best twenty per cent of college graduates. The most important feature of the scale, however, is the fact that the levels proceed by steps of approximately equal difficulty. All other intelligence tests express their results in terms of age growth and the increase of growth from one age to the next is not equal. The CAVD Scale measures increase in intelligence in levels or units of equal value at any part of the scale. At the lowest levels, A, B, C, D, and E, the test must be given as an individual test. From level F onwards it can be used as a group test. The subject is started at a level where he can achieve practically everything and is then allowed to continue up to that level where he fails practically everything. There is no time limit and so the test becomes one of power, and not of speed. The test is not easy to score and hence it will remain more of a research instrument, rather than a test to be used in everyday work in school.

31. *The Kuhlmann-Anderson Intelligence Tests.*—This is a series of tests for Grades I to XII constructed by Kuhlmann and Anderson. There are 35 sub-tests arranged in eight booklets for various grade levels. The sub-tests proceed from easy to hard and the arrangement in booklets is as follows:

Booklet for	*Tests*
Grade I	1 to 10
Grade II	5 to 14
Grade III	9 to 18

and so on. Each child is always given ten tests and his mental age on each of these ten tests is then recorded from the table

of norms. The median of these ten mental ages is then assigned as his mental age, and the I.Q. computed in the usual way. The tests have been very thoroughly standardized. A description of the method of construction of these tests has been reported by Kuhlmann (28).

32. *Trabue Mentimeters.*—These consist of a group of thirty tests gathered from various sources, including many adaptations of well-known tests as well as some original ones. In addition to intelligence tests proper, they include educational tests as well as tests for specific abilities. The whole group of thirty tests is not supposed to be given to the same individual, but the psychologist is expected to choose a series of tests suitable for each specific need. The authors indicate certain groups for certain purposes.

33. *The Myers Mental Measure.*—This group examination consists of three exercises. The whole booklet consists entirely of pictures so that no verbal responses are demanded. The directions for doing the tests are oral and, therefore, a knowledge of the English language is necessary for the examinee. The author claims that the test is suitable for first graders and that it differentiates from age 6 up to adult. Test 1 is a test of following directions; test 2, a picture completion test; test 3, a recognition of similarities. A pantomime form of this test has been constructed so that it may be given without language.

34. *The Princeton International Test.*—This is an elaborate test constructed for use in any country. The directions are given in non-language form and the subject responds by manipulating a rotator so as to indicate the response. It has been carefully prepared in order to eliminate cultural differences so far as possible, hence the elimination of the usual pencil and paper reaction. A paper form of the test requiring the usual reaction with a pencil has also been prepared. The test has been constructed by Professor C. C. Brigham of Princeton University, but so far very little has been published about it.

Combined Mental-Educational Tests.—It is not the purpose of this book to enter into a discussion or even enumeration

of the great number of educational tests that have been devised for the purpose of measuring school attainment. Only indirectly are they to be considered mental tests. Their chief function is to measure achievement in the conventional subjects studied in school. These two fields of measurement, mental and educational, have grown up and expanded together, the one seeking a more exact measure of the innate mental abilities of the individual and the other a measure of the modification of these abilities along the specific lines of school studies. The value of a combined use of these two types of measurement is at once obvious. Such a combination will lead to an evaluation of school achievement in terms of mental ability. It will help us to determine what amount of educational attainment we can expect from a pupil of a given mentality. If a child has so much innate ability, we can expect a certain degree of educational attainment in the ordinary school environment. Furthermore, such a combined use of mental and educational tests will permit us to evaluate more justly the work of the teacher and to rate more accurately the work of the school or school system.

In many educational surveys it is tacitly assumed that the mentality of the pupil-material is constant, and schools and school systems are compared solely on the basis of the results obtained by means of educational tests, whereas we know that there are great differences in the mental ability of pupil-material in different schools and in different school systems. The fact that school A rates lower than school B in arithmetic, or reading or any other school subject, cannot in itself be taken as indicative of poorer teaching or administration. Before any such judgment can be made, we need to have a mental rating of the schools. And when we have such a rating, it will often be found that the school with the lower educational score is really doing better work than its competitor, when we take into consideration the mental caliber of the pupil-material.

Most important of all, a proper use of combined mental-educational tests will enable us to prevent the enormous waste of

intelligence that prevails in our schools to-day. By finding out how much we may fairly expect of children of different degrees of intelligence, we can insist more readily on each one working up to his capacity. This will diminish the useless pressure brought to bear upon the inherently dull child, and will give the brighter child more chance to expand and utilize his powers. By doing this, we shall make the task easier and pleasanter for both the teacher and the child. The over-worked children in our schools to-day are those of lower mentality, for the system is trying to make them keep step with the average child. The laziest and most neglected are the brighter children who are kept back in order to keep pace with the average child. The more accurately we can evaluate school attainment in terms of mental ability, the more evenly can we distribute the load that each child may be expected to bear.

35. *The Pintner Mental-Educational Tests.*—These tests consist of two booklets: (1) The Non-Language Mental Tests described above for the measurement of mental ability; (2) a survey test covering the chief elementary school subjects, e.g., reading, arithmetic, grammar, history, geography. They are adapted to Grades III to VIII inclusive. The method of evaluation consists in converting by means of standard tables the pupil's score on each test into a mental and an educational index. These two indices are directly comparable and the difference between the two shows whether the child is working up to his mental capacity or not. In this way we have an extremely effective method of indicating those children who can be expected to do better school work commensurate with their mental ability. When such cases are pointed out to the teacher and superintendent, effective measures can be undertaken to correct this discrepancy between achievement and ability. The tests are well standardized and have proved very useful in several school surveys.

36. *The Illinois Examination.*—This is a combined mental educational measure with all the tests printed in one booklet. Part I is for Grades III, IV, and V. Part II is for Grades VI,

VII, and VIII. It consists of: (1) The intelligence tests, comprising tests of analogies, arithmetical problems, sentence vocabulary, substitution, verbal ingenuity, arithmetical ingenuity, and synonym-antonym. (2) The educational tests of reading and arithmetic.

The difference between the mental and educational tests is not so sharply drawn as in the Pintner Tests, because the mental tests include arithmetic and arithmetical ingenuity. All the other mental tests are language tests. The educational tests comprise only two of the school subjects, namely, reading and arithmetic.

The scores on the mental tests are converted into mental ages and intelligence quotients. This conversion is greatly facilitated by means of a table. The scores on the educational test are converted into achievement ages and achievement quotients. Separate achievement ages and quotients are given for Rate of Reading, Comprehension of Reading, and Arithmetic.

37. *The Otis Classification Test.*—This is devised for Grades IV to VIII. Part I consists of 115 educational items and Part II of 75 intelligence items. Separate ratings for intelligence and educational achievement can be calculated.

This ends our description of group intelligence tests. Thirty-seven have been described, but no claim is made that these are the only tests in existence. These thirty-seven, however, are those in common use at the present time and they represent well the different types of tests available. There are others which were once used but have now become more or less obsolete. Many tests have been described in the journals but never published. Many college tests in particular are constructed and used for college purposes and are not available for general distribution. Notable among these is the Scholastic Aptitude Test of the College Entrance Board. It has seemed wisest in this book to forego any description of these lesser known tests or of those tests not offered for sale.

In addition to tests used in the United States, there are many in use in foreign countries. Most of these would not be of

value for this country. Some, notably the British tests or non-language tests, might of course be used. No attempt will be made to describe them here. In general it may be said that they follow American models. The reader interested in such tests must look for their description in the foreign literature of our subject.

CLASSIFICATION OF GROUP TESTS

Group tests can be classified in several ways. According to content and procedure in giving directions we have the threefold division into (1) non-language; (2) non-verbal; (3) verbal. Non-language tests are entirely pictorial in nature and require no specific language on the part of the subject. The directions are all given in pantomime and by samples demonstrated by the examiner. They can therefore be given in exactly the same way to subjects speaking different languages. Non-verbal tests are those which use entirely pictorial content and therefore do not require a knowledge of reading or writing on the part of the subject. The directions, however, are given in English and therefore they can only be used with English-speaking subjects. Verbal tests assume a knowledge of reading and writing.

Non-Language Tests:

Pintner Primary Non-Language Test
Pintner Non-Language Mental Test
The Army Beta
Myers Mental Measure—Pantomime Form
Princeton International Test

Non-Verbal Tests:

Detroit First Grade Intelligence Test
Pintner-Cunningham Primary Mental Test
Pressey Primer Scale
Dearborn Intelligence Scale—Series I
Otis Primary Examination
Kingsbury Primary Group Scale

Rhode Island Intelligence Test
Detroit Advanced First Grade Intelligence Test
CAVD Scale—Levels A to E
Kuhlmann-Anderson Tests—Grade I

It will be noted that all tests for children in the kindergarten and Grade I are non-verbal tests.

Verbal Tests.—All the remaining tests are verbal. They all assume a certain degree of literacy in the subject. Some of these tests include sub-tests of the non-verbal type.

LIST OF TESTS WITH PUBLISHERS

The thirty-seven tests described in this chapter are listed below alphabetically and the publisher's name is given so that those who wish may secure sample copies in order to become more familiar with any or all of them.

	Test	*Publisher*
1.	Army Alpha	Stoelting Company, Chicago
2.	Army Beta	Stoelting Company, Chicago
3.	CAVD Scale	T. C. Bureau of Publications, N. Y.
4.	Detroit First Grade	World Book, Yonkers
5.	Detroit Advanced First Grade	World Book, Yonkers
6.	Detroit Alpha	Public School Publishing Co., Bloomington, Illinois
7.	Detroit Advanced	Public School Publishing Co., Bloomington, Illinois
8.	Dearborn, Series I	Lippincott, Philadelphia
9.	Dearborn, Series II	Lippincott, Philadelphia
10.	Haggerty, Delta I	World Book, Yonkers
11.	Haggerty, Delta II	World Book, Yonkers
12.	Illinois Examination	Public School Publishing Co., Bloomington, Illinois
13.	Kingsbury Primary	Public School Publishing Co., Bloomington, Illinois
14.	Kuhlmann-Anderson	Test Bureau, Minneapolis

Test	Publisher
15. McCall Multi-Mental	T. C. Bureau of Publications, N. Y.
16. Miller Mental Test	World Book, Yonkers
17. Myers Mental Measure	Newson & Co., New York
18. National Intelligence Test	World Book, Yonkers
19. Ohio State University Test	Ohio State University, Columbus
20. Otis Primary Test	World Book, Yonkers
21. Otis Advanced Tests	World Book, Yonkers
22. Otis Self-Administering Tests	World Book, Yonkers
23. Otis Classification Tests	World Book, Yonkers
24. Pintner-Cunningham Primary	World Book, Yonkers
25. Pintner Primary Non-Language	T. C. Bureau of Publications, N. Y.
26. Pintner Non-Language Mental	College Book Co., Columbus, Ohio
27. Pintner Non-Language and Educational	College Book Co., Columbus, Ohio
28. Pintner Rapid Survey	T. C. Bureau of Publications, N. Y.
29. Pressey Primary	Public School Publishing Co., Bloomington, Illinois
30. Pressey Cross-Out	Public School Publishing Co., Bloomington, Illinois
31. Princeton International Test	Princeton University
32. Revised Army Alpha	Psychological Corporation, N. Y.
33. Rhode Island Intelligence	Public School Publishing Co., Bloomington, Illinois
34. Terman Group Test	World Book, Yonkers
35. Thorndike Intelligence Test	T. C. Bureau of Publications, N. Y.
36. Thurstone Psychological Examination	Stoelting, Chicago
37. Trabue Mentimeters	Doubleday, Page, New York

GRADE CLASSIFICATION OF TESTS

The question as to which tests are suitable for any specific grade or school is not easily determined. This depends upon the type of school. A test that might be suitable for a second grade in one school might be far too difficult for a second grade in another school. Only general suggestions can be given below. Schools in a very backward neighborhood should choose easier tests; schools of a very select type should choose harder tests. The numbers refer to the tests in the numbered alphabetical list above.

Grade	*Tests Recommended by Authors*
Kindergarten	24, 25, 33
I	4, 8, 10, 13, 20, 24, 25, 29
II	5, 8, 10, 13, 20, 24, 25, 29
III	5, 8, 10, 11, 12, 13, 15, 20
IV	5, 9, 11, 12, 13, 15, 18, 20, 22, 23, 26, 27, 28
V	1, 2, 6, 9, 11, 12, 15, 18, 21, 22, 23, 26. 27, 28, 30, 32
VI	1, 2, 6, 9, 11, 12, 15, 18, 21, 22, 23, 26, 27, 28, 30, 32
VII	1, 2, 6, 9, 11, 12, 15, 16, 18, 21, 22, 23, 26, 27, 28, 30, 32, 34
VIII	1, 2, 6, 9, 11, 12, 15, 16, 18, 21, 22, 23, 26, 27, 28, 30, 32, 34
IX	1, 2, 6, 7, 9, 11, 16, 18, 21, 22, 32, 34
X	1, 2, 7, 9, 16, 21, 22, 32, 34
XI	1, 2, 7, 9, 16, 21, 22, 32, 34
XII	1, 2, 7, 9, 16, 19, 21, 22, 32, 34, 35, 36
College	1, 2, 7, 19, 22, 32, 35, 36
All Grades	3, 14, 17, 37

VARIOUS AIMS OF TESTING

Tests must not only be chosen to suit the appropriate level of intelligence but also to suit the aim of the worker. We may roughly differentiate three degrees of thoroughness with which testing may be undertaken.

(a) *Survey.*—Here the aim is to obtain averages of groups, rather than accurate ratings for each individual. Class comparisons and school and city comparisons are to be made. The test chosen may therefore be a brief one and it should also be one with a simple and objective system of scoring. No tests for Grades I and II are particularly well adapted for this purpose. The scoring of all of them is rather complex because of the non-verbal material. In the intermediate grades the Pintner Rapid Survey Tests are especially designed for this purpose. The Otis Self-Administering tests are also easily handled.

(b) *Classification.*—Here the aim is to get a rating for each pupil. Most of the tests can be used for this purpose. A combination of two or more tests is desirable. In such combinations the writer believes that two different types of test are better than two very similar tests. Particularly helpful is the combination of a verbal with a non-verbal or non-language test. In the intermediate grades we may well combine such tests as:

National and Pintner Non-Language
Pintner Rapid Survey A + B and Pintner Non-Language
Dearborn II and Haggerty Delta II
Otis Self-Administering and National
National and Multi-Mental
Multi-Mental and Pintner Non-Language
National and Dearborn
Otis Group and Dearborn

In the lower grades:

Pintner-Cunningham and Detroit
Pintner-Cunningham and Pintner Primary Non-Language

In the high school any two may be combined, although there is a very great similarity between all tests at this level.

(c) *Individual Diagnosis.*—Where a study of any particular child is to be made, individual tests will of course be used. In many cases, however, the results of several group tests will be very helpful. The CAVD test is particularly valuable here. For individuals of high intelligence the Thorndike Intelligence Examination is of great value. Wherever there is language difficulty or deafness or a non-English speaking background, non-language tests should always be used.

Conclusion.—This brief description of group tests must necessarily be very unsatisfactory. To convey to those unfamiliar with group intelligence tests any adequate notion of their make-up and scope, it would be necessary to reproduce dozens of the test forms, and this would be an impossibility in this book. The reader is, therefore, urged to make himself familiar with the different blanks. Many of the tests have been reproduced in psychological journals and sample copies of all of them can be bought from the various publishers. The manuals of directions accompanying the various tests give specific instructions as to how the tests are to be given and scored. They also give age and grade norms and from these I.Q.'s can be calculated in the usual way.

REFERENCES

Bird, G. E. (23). "The Rhode Island Intelligence Test." *J. of Ed. Res.*, 8, 397-403.

Bregman, E. O. (26). "On Converting Scores on the Army Alpha Examination into Percentiles of the Total Population." *Sch. and Soc.*, 22, 695-696.

Dougherty, M. L. (28). *A Comparative Study of Nine Group Tests of Intelligence for Primary Grades.* Baltimore, Johns Hopkins Press.

Haggerty, M. E. (23). "Intelligence Examination Delta 2." *J. of Ed. Psych.*, 14, 257-277.

Heckman, S. B. (24). "A Comparative Study of Group Intelligence Tests Applicable to Children of Kindergarten Age." *Contrib. to Educ.* Vol. 1, New York Society for the Experimental Study of Education, 17-44.

Kelley, T. L. (27). *Interpretation of Educational Measurements.* Yonkers.

Kingsbury, F. A. (24). *A Group Intelligence Scale for Primary Grades.* Psych. Mono., No. 153, Psychol. Rev. Co.

Kuhlmann, F. (28). "The Kuhlmann-Anderson Intelligence Tests Compared with Seven Others." *J. of Appl. Psych.,* 12, 545-594.

McCall, W. A. et al. (26). "Construction of Multi-Mental Scale." *T. C. Record,* 27, 394-415.

McGraw, M. L. and Mangold, M. C. (29). "Group Intelligence Tests in the Primary Grades." *Cath. Univ. Amer. Educ. Res. Bull.,* 4, No. 2.

Memoirs (21). The National Academy of Sciences. Vol. XV, Part I, Chap. V.

Pintner, R. (17). "The Mentality of the Dependent Child, together with a Plan for a Mental Survey of an Institution." *J. of Ed. Psych.,* 8, 220-238.

Pintner, R. (18). *The Mental Survey.* New York.

Pintner, R. (19). "A Non-Language Group Intelligence Test." *J. of Appl. Psych.,* 3, 199-214.

Pintner, R. and Marshall, H. (21). "A Combined Mental-Educational Survey." *J. of Ed. Psych.,* 12, 32-43; 82-91.

Pintner, R. (24). "Results Obtained with the Non-Language Group Test." *J. of Ed. Psych.,* 15, 473-483.

Pintner, R. (27). "The Pintner-Cunningham Primary Test." *J. of Ed. Psych.,* 18, 52-58.

Pintner, R. (27). "A New Intelligence Test for Survey Purposes." *T. C. Record,* 29, 18-23.

Poull, L. E. (27). "Clinical Values of the Rhode Island Intelligence Test and the Town Picture Game." *J. of Appl. Psych.,* 11, 68-72.

Pressey, L. W. (20). "A Group Scale of Intelligence for Use in the First Three Grades." *J. of Ed. Res.,* 1, 285-294.

Pressey, S. L. and Pressey, L. W. (19). "Cross-Out Tests." *J. of Appl. Psych.,* 3, 138-150.

Pyle, W. H. (13). *The Examination of School Children.* New York.

Sangren, P. V. (29). "Comparative Validity of Primary Intelligence Tests." *J. of Appl. Psych.,* 13, 394-412.

Thorndike, E. L. (20). "Intelligence Examinations for College Entrance." *J. of Ed. Res.,* 1, 329-337.

Thorndike, E. L. et al. (27). *The Measurement of Intelligence.* Bureau of Publications, Teachers College, Columbia University.

Thurstone, L. L. (19). "Mental Tests for College Entrance." *J. of Ed. Psych.,* 10, 129-142.

Trabue, M. R. and Stockbridge, F. P. (20). *Measure Your Mind.* New York.

Whipple, G. M. (21). "The National Intelligence Tests." *J. of Ed. Res.,* 4, 16-31.

Wood, B. D. (23). *Measurement in Higher Education.* Yonkers.

Yoakum, C. S. and Yerkes, R. M. (20). *Army Mental Tests.* New York.

PART III

THE RESULTS

CHAPTER VIII

THE APPLICATION OF INTELLIGENCE TESTS

We have sketched in Part I the early development of mental tests. In Part II we have surveyed the most important mental tests, both individual and group, which lie at the disposal of the psychologist. In Part III we shall now attempt to describe the various fields in which mental tests have been successfully used and in doing so we shall try to sum up the chief results in each field. The number of investigations in each field has, however, been so large, that it would be impossible and unwise to undertake to describe them all. Indeed, it would be unprofitable even to list them all. Only a few of the more important investigations can be described in order to bring out the main contributions that intelligence testing has made to the understanding of the different groups of individuals studied.

The beginnings of the intelligence testing movement were closely bound up with the study of mental deficiency and abnormality. It was, therefore, quite natural that much of the earliest work had to do with the selection of mentally deficient children. For a long time this phase of intelligence testing was predominant, and we have a number of studies dealing with the mental examination of feebleminded children and the examination of other children for the purpose of selecting out the feebleminded.

Closely connected with feeblemindedness is the problem of juvenile delinquency, and it was natural that the juvenile delinquent should very soon attract the attention of the psychologist. In this way intelligence testing spread rapidly from the feebleminded to the juvenile delinquent and in due time to the adult delinquent. All this work raised the pertinent question as to the relation between intelligence and crime.

It was inevitable that the study of the delinquent child should lead on quite imperceptibly to the study of the dependent child, because of the close proximity of these two types of children in our juvenile courts. We, therefore, see the psychologist comparing the mentality of delinquents and dependents and later on studying the mentality of the dependent alone, as he is found in our children's homes, orphans' homes and the like. The mere selection of the feebleminded in the school could not very long hold the interest of the psychologist. It was natural that his interest should broaden and expand into a study of the mentality of children in general. And in this broadening process the next striking phenomenon was the appearance of the very bright or superior child. School teachers and educators have, of course, always realized that some children were brighter than others, but it remained for the psychologist to be able to state definitely just how much brighter. It is doubtful whether the very decided superiority of some children was really appreciated before the advent of the psychological examination.

We now see the psychologist fairly well embarked on a study of the mentality of children in general, and a beginning made in the classification of school children into all degrees of intelligence by means of the individual scales then at his disposal. He was, however, handicapped in this work by the slowness of the procedure, the individual test taking from half an hour to an hour for each child, and the testing of whole schools requiring a corps of testers. This was rarely feasible and mental examinations would have been restricted to small groups of children for special purposes, had it not been for the coming of the group test. With the appearance of this method of intelligence examination, it became possible to test large numbers of school children. This brought in the period of school surveys on a large scale with their direct and important influence on the classification of children according to intelligence for the purpose of instruction. At the same time the group method

was extended to other groups of individuals, soldiers in the army, prisoners, college students and the like.

While the general development of intelligence tests was taking place, we note also the extension of the use of tests to special groups of individuals. Notably among these groups are the blind and the deaf. The intelligence testing of the blind and the deaf by the group as well as by the individual method has led to an increased knowledge and understanding of the specific problems encountered in the education of these children.

During all this development the use of intelligence tests in the study of racial differences has been going on. The most thorough-going studies have been made on the negro in America. Other races have only been studied slightly. The possibilities in this line of work are very great, although there are certain inherent difficulties that have not yet been fully overcome.

The field of industry and commerce is one of the latest fields in which intelligence tests have been adopted. Their value in the selection of employees and in the classification of workers with a view to assignment and promotion has now been fully recognized and they are proving an important adjunct in the solution of the problems confronting the employment manager and the business executive. In the larger sphere of vocational selection and guidance mental tests are beginning to find their place.

The remaining chapters of the book will, therefore, be devoted to the results of intelligence testing among these various groups of individuals. Beginning with the young normal child we shall progress through the grades up to college. The chapter on "The Soldier" will give us our best information about general adult intelligence. Then will follow chapters on special groups of individuals such as the feebleminded, the superior, the delinquent, dependent, deaf and blind. Racial problems will be discussed in two chapters, one dealing with the negro in America and the other with all other foreign groups. The chapter on "The Employee" will deal with the results of intelli-

gence testing in business and industry. Whether there are differences in intelligence between the sexes and to what extent intelligence is inherited are the topics for the next two chapters, and the last chapter will deal with intelligence testing in relation to miscellaneous factors that do not each deserve a special chapter.

CHAPTER IX

THE PRE-SCHOOL CHILD

The present interest in the pre-school child has aroused a similar interest in the problem of intelligence testing at this age. To be sure, even Binet in his earliest scales made some provision for tests below school age and Kuhlmann in his first revision included tests for three months, six months, and so on up. But for a long time these tests were seldom used, because of the general lack of interest on the part of psychologists in the pre-school child. Since 1920, or thereabout, the very decided interest in the infant and young child has given us many studies of intelligence at these early ages. These studies have not been primarily concerned with intelligence in a narrow sense, but rather with the general mental development of the total personality, and so we find among the tests used many that would not generally be called intelligence tests at a later stage of growth. This mixture of tests, mental and motor, is well seen in the developmental schedules of Gesell (28) quoted in Chapter VI. It is also seen in the tests of Descoeudres.

Rapid Development of the Infant.—The picture of the first year of a child's life is one of tremendous growth, not alone physically but also mentally. He makes more absolute development during his first twelve months than during any succeeding twelve-month period. This rapid development is illustrated beautifully by Gesell's comparative method of observation. This method involves the simultaneous comparison of two normal infants differing by a few months in chronological age. The simultaneous comparison of a four and a six months' infant allows one to appreciate the very marked dif-

ference in development existing between the two. And it also serves to give us some idea of the way in which the immature reactions of the four months' infant develop into the more mature reactions of the six months' baby. By thus comparing infants of many different ages, we begin to appreciate the rapidity of development during early life. Not only are we impressed with this rapidity of growth but also with what Gesell calls its orderliness. "There are," he says, "certain basic uniformities in the dynamics of development which apply to all infants, normal, abnormal, superior, inferior. There is a large system of uniformities, which characterize all normal infants and keep them traveling on highly similar routes and on highly similar time tables" (Gesell, 28, p. 124). Because of this fact it is possible to set up norms of development which show the expected normal growth.

Early Prediction.—By means of such norms of development some kind of prediction can be made for future growth. Gesell gives the results for 90 cases, 80 per cent of which were first examined when less than 18 months old and 68 per cent when less than 12 months old. These 90 cases were re-examined from time to time, some of them many times, and a final development index arrived at. If now we compare the first with the final index we obtain the following results (Gesell, 28, p. 147):*

Final Developmental Index	*No. of Cases*	*Per Cent Correctly Predicted on First Examination*
Below — 30	23	96
— 30 to — 10	19	75
— 10 to + 10	30	77
+ 10 to + 30	15	66
+ 30 and above	3	100
Total	90	81

* From Gesell, A. *Infancy and Human Growth.* By permission of the Macmillan Company, publishers.

In 81 per cent of the cases the final developmental level was correctly predicted at the time of the first examination. With reference to the very retarded cases, 96 per cent were correctly placed at the first examination. All of the three very bright cases were correctly predicted.

Pre- and Post-Mature Birth.—If an infant develops regularly, his developmental quotient will be fairly constant and Gesell finds this to be in general true. Examples of consistent regular development have been discussed in Chapter IV in dealing with the growth of intelligence. Premature and post-mature birth introduce a large error in the developmental quotients of infants. The error is particularly large during the first twelve months of life, and thereafter gradually decreases. This effect is well illustrated by Gesell (28, p. 314) in the following case of a child born two months prematurely: *

Exam.	*C.A. Months*	*Developmental Age Months*	*Corrected C.A. Months*	*Spurious Quotient*	*Corrected Quotient*
1	5	2.5	3	50	83
2	7	4	5	57	80
3	8	5	6	62	83
4	9	5.5	7	61	80
5	11	7.5	9	68	83
6	14	10	12	71	83
7	15	10.5	13	70	81
8	18	13	16	72	81
9	21	16	19	76	84
10	24	18	22	75	82
X	5 yrs.	4 yrs.	4-10 yrs.	80	83

The first column gives the number of the examination, the second shows the usual C.A. and the third the developmental age. Using the two latter columns we obtain the spurious quotients in the fifth column, spurious because they do not take

* From Gesell, A. *Infancy and Human Growth.* By permission of the Macmillan Company, publishers.

into account the two months' prematurity. These spurious quotients show a definite rise as the child grows older and they give the impression of a consistent increase in mental ability. If now, we correct the chronological age by subtracting two months as shown in column four, and calculate quotients in the usual way, we obtain the corrected quotients in the last column. These quotients show the expected stability of ordinary development. The effect of two months' prematurity is, therefore, marked during the first two years of life and does not become negligible in the calculation of the developmental or intelligence quotient until about the age of five.

In a similar manner Gesell shows the effect on the developmental quotient of a postmature birth of one or two months. If no correction in C.A. is made we find the quotient dropping from 150 at 3 months to 115 at 24 months. If a correction of one month is made, the quotient fluctuates between 110 and 120, and this is probably the best estimate. If a correction of two months is made the quotient rises from 90 to 115.

These results of Gesell are important. They point to a source of error in the I.Q. of the infant. They also suggest an additional possible cause for the greater fluctuation of I.Q.'s in the pre-school as contrasted with the older child.

Later Childhood.—Intelligence testing of children from ages two or three to ages five or six is much simpler than the testing of infants during the first two years of life. For these later ages the usual tests employed are the Kuhlmann-Binet, the Stanford-Binet and the Merrill-Palmer Scales. Table 5 shows distributions of I.Q.'s from different workers. Goodenough (28) tested 100 children each at ages 2, 3, and 4 with the Kuhlmann-Binet, Van Alstyne (29) tested 75 children from age 2½ to 3½ with the Kuhlmann-Binet. The cases reported by Reynolds (28) ranged in age from 2 to 5 and were tested on either the Kuhlmann-Binet or the Stanford-Binet. Nelson (31) tested 91 three-year-olds on both the Kuhlmann-Binet and the Merrill-Palmer Scales. Column I under Nelson in

Table 5 shows the results for the Kuhlmann-Binet, and Column II for the Merrill-Palmer Scale.

TABLE 5

Frequency Distribution of I.Q.'s of Pre-School Children

I.Q.	*Good-enough*	*Van Alstyne*	*Reynolds*	*Nelson* I	*Nelson* II
160-9	..	2	..	..	..
150-9	..	4	8	..	..
140-9	5	4	20	..	2
130-9	21	9	18	9	1
120-9	40	7	43	15	5
110-9	55	9	40	15	15
100-9	72	20	47	25	35
90-9	65	12	29	15	13
80-9	31	7	15	7	14
70-9	7	1	9	3	5
60-9	4	..	..	2	..
50-9	..	..	..	..	1
Total	300	75	229	91	91
Mean I.Q.	106.3	113.6	114.1	107.6	101.9
S.D.	16.2	21.7	19.7	16.5	14.8

We see from these distributions the very decided tendency for the mean I.Q. to be above 100. This may mean that the cases tested by these workers were in general superior children coming as they did in most cases from nursery schools. It may also mean that the standardization of the scales has been carried out on rather inferior children and hence the tests are somewhat too easy for the normal child. The first factor is undoubtedly at work in many groups of young children attending schools connected with university child-welfare institutes. Thus at Iowa, Baldwin and Stecher (27), for 105 children, ages 2 to 6, found the mean I.Q.'s on the Stanford-Binet to be as follows:

Age	*Mean I.Q.*
2	117.9
3	111.9
4	114.3
5	119.8
6	116.1

Johnson (25) also reports I.Q.'s for young children in private schools:

Age	*Mean I.Q.*
2	134.5
2½	122.3
3	114.6
3½	115.9
4	114.0
4½	115.9
5	114.4
5½	115.8

The results of Cunningham (27) allow us to compare the mean I.Q.'s on the Stanford-Binet of day nursery and private school cases:

	Mean I.Q.	
Age	*Day Nursery*	*Private School*
3-3½	99	124
3½-4	94	120
4-4½	97	131
4½-5	104	125
5-5½	94	118

These results show clearly the superior selection of the private schools, and in this respect agree fairly well with the results for private school children in later life, as we shall see in a later chapter. The I.Q. is a little higher, however, than might have been anticipated. At the same time the I.Q. for the day nursery children is also a little higher than might have been expected. These findings, therefore, point to the possibility of

the tests being somewhat too easy at these early ages. But the results in general show very decidedly the selective influence of most of our present pre-school work. The more intelligent parents appreciate the value of this work and send their children to the nursery schools and child research institutes. These children have generally good homes and superior chances for development. The children who need pre-school training most of all, namely the children of the less intelligent part of the population, seem at present not to be well represented in these schools.

Peculiar Difficulties in Testing.—In the testing of the pre-school child there are certain difficulties to be encountered which are rarely encountered with older children. Shyness, timidity, lack of cooperation, negativism and the like are more frequently encountered and, when present, are present to a greater degree than is usually the case with older children. The overcoming of such difficulties requires much experience on the part of the examiner and definite training in the testing of the young child. The implication here is not that all young children are difficult to test. On the contrary, many are extremely docile and easy to manage. A four- or five-year-old child is frequently likely to be less self-conscious than an older child and hence may make a capital subject for a demonstration test before a large audience. But, in many cases, timidity, shyness or negativism may be so great as to seriously interfere with the ordinary intelligence examination.

Negativism is the tendency on the part of the child to refuse to cooperate with the examiner. It may show itself in all degrees: mere silence and doing nothing, where it is difficult to differentiate from timidity or shyness; verbal refusals, such as "I can't," or "I won't"; active reactions against the examiner or tests, such as pushing the test away or running out of the room. Reynolds (28) has measured these negativistic tendencies in 229 children, ages 2 to 5, and scored such tendencies in 13 standard situations. She finds that negativism decreases steadily from age 2 to age 5. At age 2 refusal "is rather to be

looked upon as the usual than as the unusual thing." But from then on it should decrease sharply in the child. If it does not, the child is evidently not developing normally. The negativistic score correlates —.53 with C.A.; —.48 with M.A.; and —.09 with I.Q. The implication of all this for intelligence testing is, of course, that we must make allowance for refusals in our test procedure. The Merrill-Palmer Scale does this. In the other scales refusals are counted as failures, and thus another error is introduced into our intelligence rating.

Stability of the I.Q.—In general the I.Q. of young children has not been found to be as stable or consistent as the I.Q. of older children, although the actual evidence for this belief has nowhere been clearly set forth. If it is true, the possible explanations are numerous.

The first possible factor is a purely arithmetical one, namely, the C.A. unit used. As it is at present customary to use the same unit, namely one month, at all ages, then the younger the child the greater will be the fluctuation of I.Q. An additional month credited at one year increases the I.Q. 8 points; at two years 4 points; at 5 years 2 points; at 8 years only 1 point. We shall have to measure in finer units than months, if we wish to attain the same degree of stability as at later ages.

A second possible factor is the faulty standardization of the tests. If, as has been suggested, the tests are not as well standardized, being in general too easy, then of course the I.Q. of the child would tend to drop.

A third possible factor is the greater difficulty of obtaining a true rating, owing to the other factors discussed above under negativism and timidity. If no allowance is made for refusals, then an error in the intelligence score is introduced.

A fourth possible factor is the greater modifiability of the child at these ages. It may be that the child's intelligence is more susceptible to environmental changes at an early age than later. We have seen in Chapter IV that there is some evidence for this opinion. Some believe that the influence of attending a nursery school may affect the I.Q. There is no

such evidence, however, available. Goodenough (28) could not find any difference between the gains in I.Q. of children who had six months' nursery school experience and those who had none. Of course, a longer period might cause some real difference.

Taking all these possible factors into consideration, we can readily see why the I.Q. of the young child may not be as stable as the I.Q. of the older child.

Finally, it should be noted that all of the work reported in this chapter is based on individual testing. Very little group testing has been tried on children below the age of five. Some three- or four-year-olds who may be in kindergartens have been included in the group testing carried on there. Group tests have not been constructed and adapted for children below the regular kindergarten, and even in the kindergarten there are decided difficulties in the giving of group tests.

REFERENCES

Arlitt, A. H. (28). *Psychology of Infancy and Early Childhood.* New York.

Baldwin, B. T. and Stecher, L. I. (27). *The Psychology of the Pre-School Child.* New York.

Bridges, K. M. B. (27). "Critical Notes on Mental Tests for Children of Pre-School Age." *Ped. Sem.*, 34, 38-44.

Cunningham, K. S. (27). *The Measurement of Early Levels of Intelligence.* T. C. Contribs. to Educ., No. 259.

Descoeudres, A. (no date). *Le développement de l'enfant de deux à sept ans.* Neuchatel, Delachaux et Niestlé.

Gesell, A. (25). *The Mental Growth of the Pre-School Child.* New York.

Gesell, A. (28). *Infancy and Human Growth.* New York.

Goodenough, F. L. (28). *The Kuhlmann-Binet Tests for Children of Pre-School Age.* Institute of Child Welfare, Monograph Series, No. 2, University of Minnesota Press.

Hetzer, H. and Wolf, K. (28). "Babytests." *Zeitschr. f. Psychol.*, 107, 62-104.

Johnson, B. J. (25). *Mental Growth of Children,* New York.

Nelson, J. F. (31). *A Study of Responses other than Intellectual noted in a Simple Mental Test Situation.* To be published shortly.

Reynolds, M. M. (28). *Negativism of Pre-School Children.* T. C. Contribs. to Educ., No. 288.
Sherman, M. and I. C. (29). *The Process of Human Behavior.* New York.
Stutsman, R. (26). *Performance Tests for Children of Pre-School Age.* Genetic Psychology Monographs, Vol. I, No. 1. Clark University Press.
Van Alstyne, D. (29). *The Environment of Three Year Old Children.* T. C. Contribs. to Educ., No. 366.

CHAPTER X

THE ELEMENTARY SCHOOL PUPIL

Studies of the intelligence of the ordinary school child are very numerous, and it is not the purpose of this chapter to attempt to describe them all, but rather to indicate the different objectives of such studies in general and to give samples of the results so far obtained. Apart from the theoretical interest attaching to a knowledge of the intelligence of school children, its growth, distribution, and correlation with teachers' estimates and school work, we may say that the chief practical uses of tests up to the present time have centered around their value for the purpose of classifying children into more or less homogeneous intelligence groups, and also for predicting their future success in school work. These two purposes are intimately bound up with each other. Classification in homogeneous groups is justifiable because intelligence correlates highly with school success, and, therefore, the more homogeneous the group the more likely are the children in the group to advance together at about the same rate, be that rate relatively fast, normal, or slow. Intelligence is not by any means the only factor making for success in school, but it is a very important factor, and, therefore, the value of intelligence tests in educational administration and classroom procedure is very great.

Early Studies.—The earlier studies of school children were naturally made by means of individual tests before the advent of group intelligence tests. Most of these early studies were very narrow in scope, and they were more in the nature of trying out what the tests could accomplish. One of the earliest is that by Decroly and Degand, who in 1910 tested 45 children

in a private school in Belgium ranging in age from two to twelve years. These children belonged to a good social class and the authors found that they all tested fairly high. Decroly and Degand were inclined to suppose that Binet's norms were not satisfactory.

The testing of children to pick out the feebleminded was one of the most common purposes with the earlier workers. Here belongs the work of Irwin, who in 1913 tested 201 unselected children and found 72% normal, 12% backward and 16% feebleminded. About the same time Dougherty (13) reported the testing of 483 public school children, ages six to seventeen, giving the results in terms of numbers of years retarded or accelerated in mental age. This author is among the first to suggest that tests should be used for the purpose of classifying school children in general, although no such classification on the basis of results is reported. In 1914, the year following, Adler reports the actual classification of children based on their test records. She tested seventy first-grade children and divided them into two sections of thirty-five each on the basis of the test results. The superior section completed two terms' work in one term. In the same way eighty-nine fourth-grade children were examined and the best thirty-six cases put together in a faster-moving section, and twenty-two of these children covered two terms' work in one term. About the same time Hicks (15) discusses the value of the Binet tests for kindergarten children. Thirty kindergarten children were tested and these results were then compared with their later work in the first grade. The author concludes that the tests have high prognostic value and that they should be used for the purpose of classifying children in the kindergarten and first grade.

Dickson (20) reports the classification of children in the low first grade in a school in a poor district into three groups: (1) A border zone with I.Q.'s 85 or below; (2) a dull normal group, I.Q. from 75 to 95; (3) a normal group, plus a few who tested superior. The results of such classification are reported

to have been very satisfactory after watching the children for a year and a half. Special emphasis is laid upon the necessity for segregating the duller children.

Pintner and Noble (20) report the attempt by means of individual tests to readjust all the children in a school having Grades I to VI. After testing all the children, readjustment on the basis of mental age was attempted so far as feasible. The results of the tests were made one of the important factors in determining promotion from one grade to another. Of the 370 children examined, 22 were skipped one whole grade, 47 were skipped one half grade, 41 were retained in the same grade, 5 were demoted one half grade, one was demoted one whole grade, and the rest promoted the customary half grade. Two classes for specially bright children were formed in the 1A grade and the 2A grade, and an ungraded class was formed. A comparison of the school before and after the reclassification shows 56.7% of the pupils normally placed in grades for their mental age before reclassification, as compared with 71% rightly placed after the reclassification. Principal, teachers and superintendent reported favorably as to the progress of the children after reclassification.

Davis (22) reports the use of the Binet tests for first grade children in a city school system. The kindergarten teachers gave the tests, so that all prospective candidates for admission to the 1B grade had mental age and intelligence quotient determined. Of 277 tested first graders, 81 per cent with an M.A. of 6 and over were promoted from 1B to 1A, but only 59 per cent of those having M.A.'s between 5.8 to 6.0, and zero per cent of those below M.A. 5.8. On the basis of this type of experimentation the author says, "in the future it will be our policy to limit entrance to the first grade, in so far as feasible, to pupils who have attained a mental age of six."

The Distribution of Intelligence by Individual Tests.— There are numerous reports of individual tests of unselected school children, in addition to those we have already mentioned. Some of these studies did not aim at any practical

TABLE 6

Percentage Distribution of Intelligence Quotients

Author and Cases	Intelligence Quotient											
	40–9	50–9	60–9	70–9	80–9	90–9	100–9	110–9	120–9	130–9	140–9	150–9
Terman (19) 112 Kindergarten Children			0.9	4.5	11.6	19.6	25.0	20.5	12.5	2.6	1.8	0.9
Dickson (23) 1192 Kindergarten Children	0.3	1.2	2.5	4.6	10.3	18.0	27.2	23.6	8.8	2.7	0.5	0.1
Mitchell (22) 1113 Kindergarten and 1st Grade	0.2	2.9	8.1	15.4	23.7	20.5	18.9	7.3	3.1			
Dickson (23) 4293 1st Grade Children	0.3	1.5	4.1	10.7	19.8	25.7	22.4	11.1	3.8	0.8	0.1	
Terman (19) 149 1st Grade	0.7	4.7	8.7	18.8	20.1	14.1	16.8	8.7	4.7	1.3	1.3	
Pintner (20) 450 1st to 5th Grades			1.6	4.2	10.9	49.6		20.9	8.4	4.4		
Madsen (24) 880 1st to 8th Grades	0.3	0.2	1.8	6.4	12.4	23.1	28.5	17.4	8.1	1.7	0.2	

	56–65	66–75	76–85	86–95	96–105	106–115	116–125	126–135	136–145
Terman (17) 905 Unselected all ages	0.3	2.3	8.6	20.1	33.9	23.1	9.0	2.3	0.5
Whitcomb (22) 2360 Kindergarten and 1st Grade	1.8	4.5	10.3	22.3	27.0	22.1	8.1	2.8	0.8

purpose in school work. The results of many of these are interesting from the point of view of the general distribution of intelligence. Table 6 gives the percentage distributions, according to Binet I.Q. of various groups of unselected children. All studies do not report the results in the same steps of I.Q.; hence the two types of distribution shown in the table.

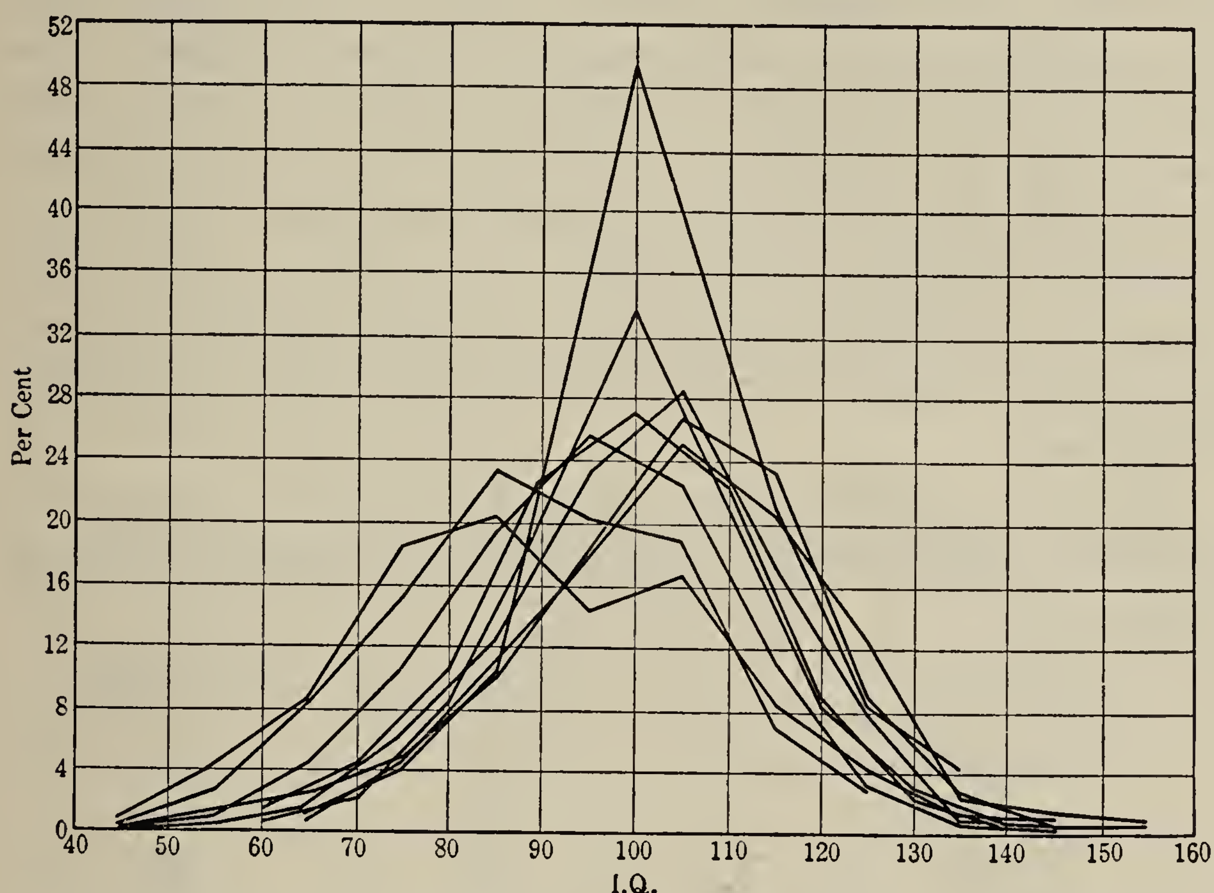


FIG. 26.—Percentage Distribution of I.Q.'s from Nine Different Studies of School Children.

Terman's (17) results of 905 unselected cases at all ages is the group upon which the standardization of his Stanford Revision is based and it is, therefore, approximately normal with the mode at I.Q. 100. It is interesting to note how the other groups deviate from this normal type and are shifted either to the upper or lower end of the distribution. Figure 26 gives the general effect of all the distributions shown in Table 6 and here we note how all the curves together give the general impression of a normal distribution, and also the shifts to the right and to the left. The two middle curves with the modes

about 100 are those of Terman and Whitcomb. The two shifted most to the left are those of Terman (19) for 149 first graders, and of Mitchell (22) also for primary children.

Any single group, grade or school may vary quite markedly from the general distribution assumed for unselected children. The selective factors of social status and grade in school are at work. The inequality between groups is strikingly shown by comparing the first three reports of primary children. The group of 112 kindergarten children tested by Terman must have been very superior, whereas those tested by Mitchell rather inferior and the first graders tested by Terman very inferior. Whitcomb's distribution at the bottom of the table shows a kindergarten group very normally distributed with percentages much like Terman's standard group for all ages combined.

Here is a distribution of Binet I.Q.s for 606 six-year-old children entering a large New York public school (Irwin and Marks. 24):

I.Q.	*No.*	*Per Cent*
0-75	83	13.8
75-85	94	15.5
85-95	149	24.6
95-105	171	28.2
105-115	77	12.7
115-125	23	3.6
125 +	9	1.7

Notice that 13.8 per cent are below I.Q. 75 as compared with 2.6 per cent of Terman's unselected cases, and 6.3 per cent of Whitcomb's cases. Again at the other end of the distribution only 1.7 per cent rise above I.Q. 125 as compared with Terman's 2.8 per cent and Whitcomb's 3.6 per cent.

The Distribution of Intelligence by Group Tests.—There are many distributions of intelligence by means of group tests. Irwin and Marks (24) give the following distribution of

I.Q.s on the N.I.T. for 1,225 pupils in Grades IV to VIII for a large New York public school:

I.Q.	*Per Cent*
49 and below	0.1
50-9	1.0
60-9	5.6
70-9	12.8
80-9	20.6
90-9	21.6
100-9	18.2
110-9	11.9
120-9	5.9
130-9	2.0
140-9	0.7

Median I.Q. = 94.3
n = 1,225

We note here that the median is somewhat below 100, and that there are very few cases of high I.Q.s.

Haggerty (23) gives a distribution of I.Q.s, by grade for 998 children in the Austin Public Schools tested by the Haggerty Delta 2 Test. Below we give the median I.Q.s by grade:

Grade	III B	III A	IV B	IV A	V B	V A	VI B	VI A
Median I.Q.	96	90	93	86	95	98	104	95

Grade	VII B	VII A	VIII B	VIII A	IX	Total
Median I.Q.	104	99	106	101	107	98.3

For the total school the median I.Q. is 98.3, but the grades vary quite markedly. There seems to be a general tendency for the I.Q. to rise with the grade. The duller pupils tend to be kept back or to drop out.

Woody (28) gives the mean I.Q.s on the Kuhlmann-Anderson Group Intelligence Test for the various grades of Grosse Ile, Mich.:—

Grade	I B	I A	II B	II A	III B	III A	IV B	IV A
I.Q.	101	102	121	110	103	103	101	100

Grade	V B	V A	VI A	VII B	VII A	VIII B	VIII A
I.Q.	97	104	101	110	95	103	102

Here the I.Q.s fluctuate considerably from grade to grade, from a low 95 in VII A to a high 121 in II B, and there is no general tendency for the I.Q. to increase with the grade.

Community Differences.—One of the earliest studies is reported by Pintner (17), who gave a set of group tests to the school population of an Ohio village. This was given as part of a general social survey of the community. The year following Paterson (18) reported the results of the same group test given to the school population of a Kansas town, also in connection with a general social survey of that community. We may, therefore, well describe these two studies together. The percentage distribution of the children as diagnosed by the tests is as follows:

Mental Rating	*Kansas*	*Ohio*
Very bright	4.2	0.7
Bright	15.6	5.8
Normal	66.0	65.6
Backward	11.4	25.3
Dull	2.4	2.6
No. of cases	332	154

The median mental index of the Ohio children is 40 (10 below normal), while that of the Kansas children is 51 (1 above normal). It is obvious that the Kansas group is superior mentally to the Ohio group. This difference in mentality is probably indicative of a like difference in the mentality of the adults in the respective communities, and it may be a determining factor in the different social conditions of the two communities. The social surveys show a decreasing population in the Ohio community and an increasing population in the Kansas community; a lack of reading in the Ohio homes, plans to increase library facilities in the Kansas town; poor recreation facilities in Ohio, a recreation director in Kansas, and so forth. In general, therefore, there seems to be a better

social condition in the Kansas town than in the Ohio village, and this healthier social condition may be ascribed in part to the superior mentality of the Kansas population. These studies are merely beginnings of what seems to the present writer a splendid opportunity for the cooperation on the part of psychology and sociology in a deeper study of community differences.

That children in different schools, whether in the same school district or in different districts, show enormous differences in general intelligence is shown abundantly by group intelligence tests. Because of the use of different group tests and different methods of rating intelligence, it is unwise to attempt a general tabular summary of these studies for comparative purposes. Results from a few sample studies will be given.

A comparison of two similar cities is made by Pressey (19), involving 1,049 cases in City A and 1,009 cases in City B, made up of children in Grade III to the first year of high school. The two cities show much the same type of pupil material although there are great differences between schools in each city. The percentage of children in City B scoring at or above the median of City A at each age is:

Age	8	9	10	11	12	13	14	15	16
Per cent ..	68	59	51	47	51	54	47	50	50

A slightly greater percentage of cases than the 10 per cent expectation in City B falls below the 10 percentile score of City A at the ages for which the results are more reliable.

Coxe (21) reports the results of the Otis test given to 24 sixth grades in 24 elementary schools in Cincinnati. The average score ranges from 61.8 to 100.1, or in terms of mental age from 10 – 8 to 13 – 10, and yet all these children are supposed to be covering the same curriculum.

Dickson and Norton (21) gave an abbreviated form of the Otis test to 1,043 eighth grade pupils in 29 schools. The median scores for these different schools ranged from 48 to

109 with a median score for all pupils of 83. Here again we see the tremendous differences in mental make-up from school to school. Individual scores range from 14 to 152 points, and yet the pupils represented by this tremendous range in ability are all in the same grade and are all supposed to be doing the same work. Comparisons with sixth and ninth grade pupils are made, and the authors say, "It is evident that ability as measured by the test in the eighth grade classes in some of the schools is equal to that in a freshman class which has survived a year in high school and is just ready to begin sophomore work; while the ability found in other eighth grades is no better than that of groups of children who have not yet completed the sixth grade. If the test means anything in terms of ability to do school work, certainly the actual accomplishments of the different classes graduating in January, 1920, must have varied widely."

A survey of all the schools in a small city by means of several intelligence tests was made by Holley (20). In all, 2,030 children in all grades in the eight elementary schools and the one high school in the city were tested. Although there are differences in the median scores for the same grade from school to school, the differences are not very striking. Evidently the mental make-up of the various sections of this city is fairly similar on the average. Again a study of the tables showing the distribution of the I.Q.s, as calculated from several group tests combined, for each school and each grade separately, shows great similarity from school to school. There are, however, great variations among the grades. The highest median I.Q. for any one grade is 117 and the lowest 93. The percentage distribution of I.Q.s for the total 2,030 cases is:

I.Q.	*Per Cent Pupils*
150-9	0.2
140-9	1.0
130-9	4.0

I.Q.	*Per Cent Pupils*
120-9	11.4
110-9	22.9
100-9	26.3
90-9	19.7
80-9	8.9
70-9	4.0
60-9	1.4
50-9	0.5

The median I.Q. is 106. The distribution is fairly symmetrical with a slight skewness towards the higher I.Q.s. The town is a university town and this may account for the higher percentage of brighter pupils as contrasted with the backward group.

An extensive survey of the schools of Utah is reported by Snoddy and Hyde (21) in which more than 15,000 children were tested. The differences in mental ability of different schools or communities are not given and the authors content themselves with total distributions of those taking the elementary and those taking the advanced examinations respectively. The two curves for these groups show very close approximations to the normal probability distribution.

Results by Woody (24) show the relationship between intelligence and size of city. He presents the mean scores for the National Intelligence Test given in three groups of cities, comprising 23 cities in all:

	Population of Cities			
Grade	*Less Than 5,000*	*5,000 to 10,000*	*More Than 10,000*	*All Cities*
III	34.0	36.7	40.2	38.6
IV	58.0	66.3	68.4	66.1
V	79.3	89.0	90.0	88.0
VI	98.1	106.1	109.6	107.3
VII	117.2	123.1	126.0	124.4
VIII	133.1	142.2	141.6	140.4

We note that in every grade the score increases with the increase in population. The scores of the largest cities are about 10 points higher than those of the smallest. Does this show the selective influence of the large city in attracting the more intelligent people?

Similarly Coxe (25) shows that villages differ in the mental ability of their school children. Here are the median mental ages of the Otis Classification Test for several villages for Grades V and VI:

Village	*Grade V*	*Grade VI*
1	10.3	10.9
2	10.6	11.7
4	11.3	12.7
8	11.6	12.9
9	10.0	10.9
12	10.8	11.9
14	10.1	10.4
23	10.7	12.2
25	10.1	11.7
26	10.5	11.6
27	11.2	12.3
28	10.6	12.0
31	9.7	11.2
33	10.3	10.9
35	10.8	11.5
36	10.6	12.3
37	11.3	13.2
Median	10.6	11.7

There is a range of almost 2 years in Grade V and of almost 3 years in Grade VI.

Kempf and Collins (29) in a mental survey of two Illinois counties find the northern county decidedly superior to the southern county in intelligence. For 4,483 children in the northern county the median I.Q. is 102; whereas for 677 children in the southern county the median I.Q. is only 88.

Comparison of City and Country Children.—Book (18) uses the results of a survey of 1,165 children in several schools in Grades III to VIII to draw a comparison between city and country children. The median scores for each age for the two groups are as follows:

Age	*City*	*Country*
8	67	50
9	74	58
10	93	62
11	105	89
12	116	97
13	125	107
14	122	110
15	115	117

Thus at each age, with the exception of age 15, the elementary school children of the city show higher intelligence scores than those of the country. The author remarks, "If we take the years 8 to 13 inclusive for the city and country we find that only 24 per cent of the country children rate above the medians for the city and only 2 per cent score above the highest 10 percentile of the city children. Of the country children 27 per cent rate below the lowest 10 percentile for the city." Pintner (18) compared four city schools, a village school and a rural school. The median percentiles of the four city schools were 58.5; 58.5; 47; 44.5; for the village school, 30; and for the rural school, 17. Differences in the actual native ability of the various groups examined and not differences in the methods of teaching and the like are presumed to be the reason for this variation. Pressey and Thomas (19) also find the country children below the city children. They compare 321 country children with the age norms for city children and find the following percentages testing at or above the median for the city children:

Age	10	11	12	13
Per cent	29	33	21	25

These percentages are very similar to the percentage found by Book mentioned above. Pressey and Thomas furthermore divide their country children into those living in a poor district where the land is hilly and the soil poor and those living in a good district. They find the children in the better district to be superior to those in the poorer district. In the former district 36 per cent are above the median for the city children as contrasted with only 20 per cent in the latter. This would fit in with the hypothesis that the more intelligent tend to find the better land and better opportunities.

Van Wagenen (29) in an extensive survey of about 4,000 rural children and 8,000 town and city children in Grades VII and VIII finds that "the seventh grade rural pupils are approximately four months younger mentally than the town and city pupils of the seventh grade, while the eighth grade rural pupils are approximately six months younger mentally than the town and city pupils of the same grade."

In the comprehensive survey by Kempf and Collins (29) of two Illinois counties we find median I.Q.s for urban and rural children as follows:

	Northern County	*Southern County*
Urban	103.5	91
Rural	95	84

Considering native whites alone, we find in the northern county a median I.Q. of 108 for urban children and 99 for rural children.

In England Thomson (21) also comments upon the selective influence of cities but he also found a large percentage of high I.Q.s in small schools in remote country districts far away from large cities:—"The distribution of intelligence suggested by the tests is such that the highest ability appears to be found close to the cities and far away from the cities, the intermediate areas having few cases of high ability, as though they were drained by selection." This work was done in the

county of Northumberland, England, and the superior intellect found in these remote country schools may be due to very special causes. It does not seem to be found in other remote country schools in Yorkshire or the southwest of England (Marsden, 29) nor have any workers in the United States reported similar findings in remote country districts.

In general, therefore, it would appear as if the urban districts rate higher in intelligence than rural districts and that this is due to the migration of superior intelligence to the cities. This does not mean that all rural districts are low in intelligence. On the contrary some may be very high, and the difference in any case is not very great. Some of the difference that exists may be due to the fact that the tests themselves favor city children or that city children are more familiar with tests, but all of the difference is not likely to be explained away by these suggestions.

Private Schools.—Several surveys of the intelligence of children in private schools in the East have been made and all of these agree in showing that the pupils of such schools are very markedly above the average in intelligence. Rogers (28) reports results from schools in the neighborhood of Philadelphia, Wood (29) from schools in New York and New England and Dearborn and Cattell (30) from Boston. Table 7 shows the percentage distribution of these I.Q.s for private children. In the last column for purposes of comparison is given the percentage distribution of I.Q.s of 4,925 public school children mainly in New York City. The median I.Q. for the private school children is 115 to 119, and for the public school slightly below 100. If we consider those with I.Q.s above 120 as bright, then we have in the private schools from 35 to 41 per cent in this group as contrasted with about 10 per cent in the public schools. Conversely, we find a very small percentage of dull children below 80 I.Q. in the private schools, only 1.1 per cent of the Rogers data and 0.1 per cent of the Wood data; whereas the public schools show about 19 per cent.

TABLE 7

Percentage Distribution of I.Q.s in Private and Public Schools

I.Q.s	*Private Rogers*	*Private Wood*	*Private Dearborn*	*New York Public Schools Pintner*
150-up	2.4	0.6	1.9	0.5
140-9	4.0	1.0	4.4	0.7
130-0	10.0	8.0	14.4	2.5
120-9	19.0	31.6	24.8	6.6
110-9	29.0	35.8	26.9	12.6
100-9	23.0	16.9	16.2	19.8
90-9	8.0	4.9	8.0	20.6
80-9	3.0	1.1	2.6	17.2
70-9	0.7	0.1	0.5	12.7
70 and below	0.4	...	0.2	6.6
Median I.Q.	115	118	119	96
n	3,522	1,157	1,295	4,925

Another way of showing this difference in mental calibre is to compare the average mental age by grade. Rogers gives these averages for private school children as contrasted with Terman's averages:

Grade	*Private*	*Public*
I	7-10	6-10
II	8-10	7-11
III	9-11	9-0
IV	10-9	9-11
V	12-0	11-0
VI	13-5	12-1
VII	15-0	13-1
VIII	15-10	14-2
IX	16-6	15-4

This big gap in intelligence between public and private school children seems to be the customary thing at least in the

eastern part of the United States where private schools are numerous. Such schools are primarily training grounds for the élite in intellect. Their problems must necessarily be very different from those of the average public school. Their methods of instruction, courses of study and general educational policies may differ radically from those of the public schools. Educational innovations that are successful in private schools may not be at all suitable for public schools because of the decided difference in intelligence between the pupils of these two types of school.

Overlapping in Ability.—Madsen's (22) tests of 12,000 children bring out strongly the facts as regards the overlapping of ability among children in different grades. He also uses the data to show the well-known fact that the older children in a given grade are the duller and the younger are the brighter. He shows for each grade facts similar to those which we give below for grade IV:

Average I.Q. at Age

8	*9*	*10*	*11*	*12*	*13*	*14*
142	107	86	62	56	59	46

In this grade the youngest children have an average I.Q. of 142 and the oldest of 46. There is, of course, no criticism necessarily implied in this sort of distribution, provided the mental ages of the children are more or less equal. In the above distribution this is not the case. Turning the data into mental ages and considering age 8 as equivalent to 8 yrs. 6 mos., and so on for the other ages, we get the following facts for grade IV:

Average M.A. at Age

8	*9*	*10*	*11*	*12*	*13*	*14*
12-1	10-2	9-0	7-2	7-0	7-11	6-8

Here again we see that the younger children of C.A. 8, 9 and 10 are in absolute intelligence above the older children of ages 11, 12, 13 and 14. If, as Terman says, a mental age of

10 should be the standard for grade IV, we might expect the following I.Q.s for children of different chronological ages:

Average I.Q. for Age

8	*9*	*10*	*11*	*12*	*13*	*14*
118	105	95	87	80	74	69

A comparison of this ideal arrangement with the one actually found by Madsen shows much less difference in the spread of I.Q.s for the different chronological ages.

Terman (22) discusses this problem of overlapping or the heterogeneity in intelligence within a given grade and shows many sample distributions. He says, "The condition may be summed up by the statement that, in general, from 20 to 25 per cent of the pupils of a given grade have attained a mental age about as high as the median mental age of the next higher grade, while the lowest 20 to 25 per cent in the same grade are about as low in mental age as the median for the next grade below."

The great differences in intelligence that may be expected in the same grade in different schools or in different rooms of the same school are well shown in Terman's (20) report of the intelligence of five first grade classes:

Room	*Median M.A.*	*Median I.Q.*
1	5.7	76
2	6.0	87
3	6.0	85
4	7.2	108
5	7.8	112

Here we range from a group whose median I.Q. denotes border-line mentality and in which there are probably many feeble-minded children, to a group whose median I.Q. denotes superior intelligence and which, therefore, contains more than fifty per cent superior children.

Mental Age Standards for Grading.—In the placing of school children in grades in terms of their mental age, what

mental age standards should be adopted? On the assumption that the average age of school entrance is about 6½ years, Terman (19) gives the following standards:

Grade	Mental Age
I	6½ to 7½ or about 7
II	7½ to 8½ or about 8
III	8½ to 9½ or about 9
IV	9½ to 10½ or about 10
V	10½ to 11½ or about 11
VI	11½ to 12½ or about 12
VII	12½ to 13½ or about 13
VIII	13½ to 14½ or about 14
High School I	14½ to 15½ or about 15

The median mental ages actually found according to Terman (19) are as follows:

Grade	Median M.A.	No. of Cases
I	6-10	341
II	7-11	189
III	9-0	181
IV	9-11	253
V	11-0	226
VI	12-1	236
VII	13-1	193
VIII	14-2	180
High School I	15-4	137

CLASSIFICATION

Intelligence tests are frequently used in school work in order to help in the formation of more homogeneous groups. Sometimes this ability-grouping is based entirely upon intelligence tests, but generally other factors, such as school achievement, general health and teacher judgment, are also taken into account. So far as intelligence measures are concerned, we have in the M.A. and I.Q. two different indicators. The M.A.

shows the present attainment of the pupil and indicates the grade in which he should be able to work. The I.Q. is a measure of his probable rate of progress and indicates the section of the grade (slow, average or fast) for which he is best suited.

Mental Age to Determine Grade.—In placing a child, therefore, the M.A. is first considered in order to adjust him as well as possible to the grade in which he may do his best work. If he is a very bright child, it may be feasible to advance him a grade or so above that in which he now is, or at least let him move more rapidly through such grades. Skipping a grade or half grade may be very desirable, because many bright children are wasting their time and becoming bored and are doing poor work in grades where the work is far too easy for them. Skipping or acceleration may at times be advisable with a child of high M.A. even when the educational age or school mark is relatively low. Such treatment may be a valuable incentive toward work more in accordance with his real intellectual capacity. At the same time no acceleration or skipping should be attempted, if the child is in poor health, or if the teachers concerned are not favorably disposed toward the experiment. A teacher must cooperate heartily in helping the child make the necessary new adjustments.

Proper classification is very important in the kindergarten and first grade. The first year or two of school life may frequently determine the amount of effort a child will put into his future school work. A very bright child with mental ability well beyond that required for first grade work, if left in the first grade, will not need to exert any mental effort to accomplish the work. He receives commendation from the teacher for doing easy tasks and so learns to be satisfied with that type of performance. And so he may continue during his scholastic career, satisfied with work far below the level of his ability.

In this general way, therefore, the M.A. will help to deter-

mine the proper grade. Bright children can be accelerated. Dull children can be prevented from being pushed ahead too fast, so that their school work becomes a burden to them and their only wish is to leave school as soon as possible. The greatest satisfaction is achieved when children are working at their true level of ability.

The I.Q. for Sectioning.—In large schools, where there are several sections of each grade, a further refinement in classification can be made. Supposing we have placed all our children in the most appropriate grades according to their mental ages and a due consideration of any other important factors, then we will find that the children in any one grade will vary greatly in I.Q. There will probably be a scattering of I.Q.s all the way from about 80 to about 150. The actual range of I.Q.s in any particular case will be determined by the neighborhood in which the school is located. Here is the frequency distribution of the I.Q.s of a V-A grade tested on the National Intelligence Test:

I.Q.	*Frequency*
130-139	2
120-129	2
110-119	14
100-109	16
90-99	22
80-89	18
70-79	11
60-69	7
50-59	1
Total	93

Now, even if all these children have the same mental age at the beginning of the year, and therefore are capable of doing the same work, yet the various I.Q.s indicate various rates of growth, various rates of ability to progress with their school work. To the extent that the I.Q. is constant, the child

with an I.Q. of 50 will only grow half a year in mentality during the year; the child with an I.Q. of 100 will grow a whole year; and the child with the highest I.Q. will grow a year and a third. If, therefore, we pay no attention to their I.Q.s in classification, we shall have increasing differences in mental age as the year progresses. The natural thing to do, therefore, is to divide them into sections according to their I.Q.s, so that the classes will be as homogeneous as possible from the point of view of rate of growth of mental ability. In the example given above, three rooms and three teachers were available. The thirty highest I.Q.s were put in one section, the next thirty-five in another section and the lowest twenty-eight in another section. Just where the lines should be drawn for the different sections will always depend upon the number of rooms, the size of rooms, the number of teachers, and so on. In general it is better to make the highest and lowest sections somewhat smaller than the intermediate sections, if possible, because the differences in I.Q. will be greater in the former than in the latter.

The general scheme of classification outlined above does not take adequate care of extreme cases. Children of extremely high intellect would be so far advanced beyond their mates of similar chronological age, and children of extremely low intellect would be so far retarded, that difficulties of adjustment would undoubtedly take place. A seven-year-old boy with an M.A. of 11-7 and an I.Q. of 165 in the sixth grade or a thirteen-year-old boy with an M.A. 7-7 and an I.Q. of 58 in the second grade would be examples of socially maladjusted children. For such extreme cases special classes are desirable.

Other Classification Methods.—The general plan of using the M.A. for grading and the I.Q. for sectioning takes no special account of educational attainment. McCall and Bixler (28) recommend a method which pays more attention to educational attainment on standard tests. Having chosen a battery of educational and intelligence tests the *G* or grade

scores for each test are computed. Then the various *G* scores for the several educational tests are combined into a single *G* score for educational attainment. The *Gi* (intelligence) and the *Ge* (education) are now combined, by weighting if desirable, into a final *Gp* or grade score for placement. This *Gp* will determine the grade a child should be in, considering his intelligence and educational attainment. Very complete tables have been constructed for the interpretation of these *G* scores, taking into account the month of schooling in which the testing takes place and the type of school, i.e., whether it attempts 0.9 or 1.0 or 1.1 standard grades per year.

Classification in Practical Use.—The work of Brooks (22) may be mentioned as a good example of the use of group intelligence tests in a school district by a practical school man in the field. His distribution of 592 children according to grade and ideal mental age (allowing two years as normal for each grade, e.g., M.A. 6 to 7 – 11 for grade I, 7 to 8 – 11 for grade II, etc.) shows 89 cases or 15 per cent of the pupils working "above or below their mental capacities. Of these 7.1 per cent are working above or trying to, and 7.9 per cent are working below." He then goes on to show how to improve the grouping of the children, taking into consideration the practical necessities of the school situation. He shows a decided preference for the intelligence test as contrasted with the achievement test for classification purposes.

The educational achievement of children in grade VIIB, sectioned on the basis of intelligence tests, is reported by Theisen (22). After one semester's work the classes were tested by means of standard educational tests and the brighter sections were shown very definitely to have covered much more ground than the others. Most interesting in this report are the comments of the teachers upon this method of sectioning classes.

That homogeneous classification reduces the number of failures is pointed out by Torgerson (26). The percentage of failures among 2,500 pupils in the first six grades was reduced

from 12.2 per cent before classification to 7 per cent after classification; and among 686 pupils in Grades VII to IX the reduction was from 16 to 13 per cent. Other workers report similar reductions in failures after a period of homogeneous classification.

It is also claimed by some that ability grouping leads to increased educational attainment. Thus Torgerson (26) compared his 1,222 classified pupils with 4,327 other non-classified pupils on the Thorndike-McCall Reading Test. His classified pupils had 82.8 per cent normal grade reading ability as contrasted with 61.6 per cent of the unclassified. Furthermore only 7.2 per cent of the classified were retarded more than one year in reading, whereas 23.9 per cent of the unclassified were so retarded. There is no direct evidence, however, that this superiority in reading was caused only by ability grouping.

Some results of the Detroit plan of classifying children into three groups, X (high), Y (medium) and Z (low), have been reported by Berry (22). Pupils are tested by the Detroit First Grade Intelligence Test on their entrance to the first grade. For 10,316 pupils at the end of the first semester there were the following percentages of promotions: 96.7 for the X group, 85.2 for the Y, and 62.6 for the Z. Eleven times as large a percentage of Z pupils failed of promotion as of X pupils. Similarly in the amount of work done during a semester, we find 91.5 per cent of the X group doing one semester's work or more, 71.5 per cent of the Y group and only 45.1 per cent of the Z group. More than half of the Z group failed to complete one semester's work in one semester. Keener (26) reports the results of classifying 218 VIIB pupils into six sections on the basis of Otis Classification scores. This sectioning was done in September and followed by the Stanford Achievement Test in January and May. The mean educational achievement scores for each section increase steadily from the lowest to the highest section, and the improvement in educational score from January to May shows least improvement in the three lowest sections and most in the three highest

sections. There are other studies of the same type as the last two mentioned. The defect of such studies is clear. They lack control groups. They do not show that the gains made are due to homogeneous classification. Similar gains may be made among heterogeneous groups.

Worlton (27-28) attempts to supply the above-mentioned deficiency and compares bright pupils in homogeneous sections with bright pupils in heterogeneous sections. He finds that the educational achievement scores of the former are appreciably higher than the educational scores of the latter group and he attributes this difference to the method of classification. Other investigators, however, have not been able to find any improvement in achievement resulting from homogeneous grouping. The most careful study leading to this negative conclusion is that of Purdom (29). Although dealing with first year high school pupils, we will mention it in this chapter. His sectioning was based on the Terman Group Test. Sections were equated for age, sex and intelligence score, some being homogeneous and some heterogeneous. He found that homogeneous sections did not gain more on standard tests or on semester grades. No degree of intelligence, whether bright, normal or dull, showed the advantage of homogeneous grouping. The teachers believed that homogeneous grouping was better, but Purdom could find no evidence of this in the achievement of the pupils.

From such studies as these we may conclude that the classification of pupils into homogeneous sections is no universal panacea for all educational ailments. Mere classification may give more opportunity for individual differences in intellect to assert themselves, but this does not seem always to follow. Evidently much still remains to be done by the teacher. Courses of study and methods of instruction must be adapted to the different levels of ability and much experimentation is still required in order to find out how this may best be accomplished.

Combined Mental-Educational Tests.—The methods proposed by various workers for combining intelligence and educational tests in school measurement have been discussed in Chapter V and the special sets of tests devised for such methods have been described in Chapter VII. We shall give here sample results from such combined measures.

Pintner and Marshall (21), using the scheme of mental and educational indices, report results from 56 schools of various sizes, both city, village and rural. They show how the combination of mental and educational tests may be used to pick out the children who are not working up to mental capacity, and in the same way classes and entire schools not making use of all the intelligence possessed by their pupil material may be discovered. They go on to point out that this double measure is the only fair way to make comparisons as to the relative efficiency of different schools or school systems. The usual comparison between schools as to gross educational accomplishment is not a fair measure of efficiency. They show how schools vary extraordinarily in the extent to which they are utilizing the intelligence of their pupils. In general, children possessing superior intelligence are the ones who are not working up to possible accomplishment, and the final verdict is that our educational system is failing to make use of vast stores of intelligence which lie hidden and undiscovered.

Murdoch (22) reports the Achievement Quotients of all pupils in grades III to VIII in a private school. Separate A.Q.s for reading and arithmetic were calculated. Because very high and very low scores on our educational tests do not have equivalent educational ages, the author was forced to make hypothetical extensions of the scales at both ends. This is the difficulty and drawback of the educational age-mental age method upon which the A.Q. depends. The author's results show most of the class A.Q.s in reading above 100, and in arithmetic 4 above 100 and 6 below. The A.Q.s were found useful in helping to reclassify the children. "Often it [the A.Q] was particularly useful in indicating that some pupil

already at the 95 or 100 percentile for his grade should be advanced to a higher grade, because his A.Q., in spite of his high grade standing, was low." She further finds that in the case of many gifted children more accomplishment should be demanded because their A.Q.s were far below 100. The A.Q. also indicated children who required special testing, and lastly it was used to evaluate the teaching efficiency of the teachers.

That the classification of pupils into homogeneous groups helps to raise the A.Q.s is pointed out by Torgerson (22). A sixth grade showed an increase in reading ability from an A.Q. of 114 to 162 during one semester when pupils were classified according to intelligence, as contrasted with an increase from 100 to 114 the previous semester when they were not so classified.

The Achievement or Accomplishment Quotient is stressed by Stebbins and Pechstein (22) as being the best measure of the efficiency of the teacher. Children in grades IV to VII were tested by means of an intelligence test and several educational tests. The median A.Q.s for the various grades show that in general those made up of the duller pupils receive the higher A.Q.s.

In Van Wagenen's (26) work on educational diagnosis the educational test results are always reported in terms of the mental age of the child. Danielson (29), working with superior children of I.Q. 125 and above, found their A.Q.s below 100. This is usually the case because they are generally in classes below their M.A. and so they rarely are able to achieve in accordance with their high M.A.'s. Danielson, however, gave these children a wide course of reading in rich and varied materials for two years, and at the end of that time found that the A.Q.s as based on the Stanford Achievement Tests were slightly above 100.

RELATION OF INTELLIGENCE TO ACHIEVEMENT

Almost all the studies mentioned in this chapter and very many others discuss the relationship between the results obtained on the intelligence tests and educational achievement, be the latter measured by objective educational tests or by school marks or by teachers' estimates of ability. A detailed discussion of the literature on this subject would require almost a book in itself. We can only give here a random sampling of the results obtained and refer the reader also to Chapter V, where correlations have been given with reference to our discussion of the validity of intelligence tests.

One of the most detailed and elaborate studies is reported by Gates (22). School achievement was measured by means of three objective educational tests in grade I to about 16 in grades III to VIII. Intelligence was measured by the Stanford-Binet and by seven to ten group intelligence tests, both verbal and non-verbal. The correlations of the educational tests with the three types of intelligence tests are as follows:

Grade	*Binet*	*Verbal Group*	*Non-Verbal Group*
I	.36	..	.30
II	.44	..	.23
III	.47	.65	.22
IV	.42	.54	.22
V	.51	.49	.17
VI	.67	.57	.29
VII	..	.52	.08
VIII	..	.47	− .15

There were about 20 pupils in each grade.

Burt (21) does not give his results for each grade separately. He reports a correlation of .91 between the Binet tests and school work for 689 children ranging in age from 7 to 14. The partial coefficient, with age constant, is .68.

Burt's correlation is extremely high. Few workers have

found correlations above .60. Most of the correlations reported are more like those of Gates shown above. Kempf and Collins (29) correlated I.Q.s and school marks for 3,747 children of all grades and found a correlation of .39.

Some examples of correlations between intelligence tests and specific subjects may be quoted. Gates (22) found the following correlations between Binet tests and certain subjects measured by educational tests in three grades as follows:

Grade	*Reading*	*Arithmetic*	*Spelling*
IV	.36	.35	.11
V	.41	.25	.37
VI	.69	.30	.45

For a large group of children, age 7 to 14, Burt (21) reports correlations between the Binet and school marks as follows:

Composition	.63
Reading	.54
Dictation	.52
Arithmetic (Problems)	.55
Arithmetic (Mechanical)	.41
Writing	.21
Drawing	.15
Handwork	.18

Hollingworth (18) in a study of spelling reports correlations of .16 to .47 for various groups of children between Binet tests and spelling ability.

We note, therefore, in practically every case a positive correlation between educational achievement and intelligence. The coefficients range from very small ones to very high ones. Coefficients between .30 and .60 seem to be most frequent. The great variation in these coefficients is due to many factors, such as, the homogeneity or heterogeneity of the group, the reliability of the educational ratings, the great differences in the intelligence tests used, the range of ages included in any one group, the thoroughness with which education and intelli-

gence have been measured in each case, and other similar factors. In general, however, the results clearly indicate that the thing measured by our intelligence tests is one of the factors making for success in school work.

School Failures.—As a consequence of this positive correlation between intelligence and school achievement, we find that those who fail to be promoted have lower I.Q.s in general than those who are promoted. Thus Stalnaker and Roller (27) studied 100 non-promoted children and found 91 per cent with I.Q.s below 90 and 31 per cent with I.Q.s below 70. Maddocks (27) studied 100 children who had failed in a subject and found 56 per cent with I.Q.s below 90. In a more extensive survey of children two or more years over age for their grade in New York City, O'Flaherty (24) found 3,826 cases in fourteen schools or about 12 per cent of the total enrollment. Intelligence tests were given to 3,290 cases and a distribution of the I.Q.s shows:

I.Q.	*Per Cent*
Below 70	21
71-90	57
91-110	21
111 and above	1

Only 22 per cent of these over-age children are of normal intelligence; 78 per cent are dull or backward.

Conclusions.—It is evident from the survey in this chapter of the use of intelligence tests with school children that such tests are at present being used very widely and for many different purposes. They have not yet come to be considered an integral part of every school system, but the time is not far distant when they will be considered as essential for the health, happiness and advancement of every child, just as a good school system now considers as necessary, physical examinations, adequate physical exercise, suitable buildings and equipment, adequately trained teachers, and the like.

At present tests are being used for the purpose of classifying

children into more or less homogeneous groups. Often this results in the formation of separate fast or slow sections, not in the sense of extra sections for the deficient and the very superior, but simply as being relatively quicker or slower than the middle or normal group. Most reports of such classification are favorable, and justice is being done to all types of children by this procedure. The teacher is helped, for her work becomes more uniform and even. She is not continually distracted by the maladjusted child. Discipline is greatly improved, for sufficient interesting work will keep the child out of mischief. A natural corollary to this movement of sectioning according to ability will be the gradual differentiation of the course of study in terms of the capacities of the various groups of pupils. Little has been done in this respect, but the indications are that attempts in this direction will be made now that the individual differences of pupils have been so clearly demonstrated.

The results of intelligence tests are also being used to help determine the promotion, acceleration or demotion of pupils. The re-adjustment of a whole school or the proper adjustment of an individual is greatly helped by an adequate intelligence measure. Children may often be stimulated by promotion or acceleration, even though their school work is none too good, if they possess the requisite amount of intelligence. We have noted again and again the reports of bright children not working up to capacity and the reaction that often occurs when they are confronted with tasks that challenge their ability. Demotion of the slow who have been pushed along by a mechanical system of promotion is seldom reported by workers and it seems to be regarded as of questionable value. If so, it means that in future we must be more critical of our methods of promotion and not allow the dull child to move on to higher grades simply because of his chronological age, and by so doing to become a problem and a burden to the teachers in those grades.

A third use of intelligence tests is in the fields of educa-

tional and vocational guidance. All are agreed that intelligence tests are simply one item in helping to give educational and vocational advice. They can, however, be made to function admirably if judiciously used. Because a child has a certain I.Q., there is no reason to believe he will follow with profit a certain type of training or make good at a certain job. Interest, desire, will-to-succeed, previous training and many other factors enter into the problem. But, unless he has a certain I.Q., we may now say, with a fair degree of certainty, that his chances for success in this educational course or in that vocation are very small or zero; and, conversely, if he is above a certain I.Q., we may encourage him to go on with certain studies or to pursue certain ambitions.

A fourth use of intelligence tests, particularly when combined with educational tests, is the adequate rating of schools and teachers. In addition, their possibility for adequately measuring the value of different methods of teaching is indicated. We now know that the quality of the pupil-material must be taken into account in measuring the efficiency of a school or teacher or the value of a teaching method. Undoubtedly the near future will see much work with combinations of intelligence and educational tests.

Apart from these practical uses of intelligence tests in the schools, this chapter has described the importance of test results in indicating the individual differences among children. The presence of such differences has, of course, always been more or less recognized. We are now, however, measuring them accurately for the first time, and the differences are greater than was ever suspected. Schools and classes differ radically in the intelligence calibre of their pupils and so do communities. Light has been shed upon the much debated question of the relative intelligence of city and country children, and, so far the verdict seems to be in favor of the city child. But the work in this field is only beginning and the future should see important studies in community differences. Geographical, economic and industrial conditions are not the

only reasons for the success or failure of communities, towns or cities. The mental make-up of the inhabitants will surely be shown to be of great and far-reaching importance.

REFERENCES

Adler, M. (14). "Mental Tests used as a Basis for the Classification of School Children." *J. of Ed. Psych.*, 5, 22-28.

Berry, C. S. (22). "The Classification by Tests of Intelligence of Ten Thousand First-Grade Pupils." *J. of Educ. Res.*, 6, 185-203.

Book, W. F. (18). "Variations in Mental Ability and its Distribution Among the School Population of an Indiana County." *Bull.*, the Extension Division, Indiana Univ., Vol. IV, No. 4, 100-131.

Brooks, S. S. (22). *Improving Schools by Standardized Tests.* Boston.

Burt, C. (21). *Mental and Scholastic Tests.* London County Council, London, England.

Coxe, W. W. (21). "School Variation in General Intelligence." *J. of Ed. Res.*, 4, 187-194.

Coxe, W. W. (25). "A Study of Pupil Classification in the Villages of New York State." *Bull.*, Univ. of the State of N. Y., No. 841.

Danielson, C. L. (29). "A Study of the Effect of a Definite Course of Reading in General Literature upon Achievement in Content Subjects with Children of Superior Mental Ability." *J. of Ed. Psych.*, 20, 610-621.

Davis, H. (22). "Intelligence Tests in Public Schools in Jackson." Chapter III, *Twenty-first Yearbook N. S. S. E.*, 131-142.

Dearborn, W. F. and Cattell, P. (30). "The Intelligence and Achievement of Private School Pupils." *J. of Ed. Psych.*, 21, 197-211.

Decroly, O. et Degand, J. (10). "La Mesure de l'intelligence chez les enfants normaux d'apres les tests de MM. Binet et Simon." *Arch. de Psych.*, 9, 81-108.

Dickson, V. E. (20). "What First Grade Children Can Do in School as Related to What Is Shown by Mental Tests." *J. of Ed. Res.*, 2, 475-480.

Dickson, V. E. (23). *Mental Tests and the Classroom Teacher.* Yonkers.

Dickson, V. E. and Norton, J. K. (21). "The Otis Group Intelligence Scale Applied to the Elementary School Graduating Classes of Oakland, California." *J. of Ed. Res.*, 3, 106-115.

Dougherty, M. L. (13). "Report on the Binet-Simon Tests given to 483 Children in the Public Schools of Kansas City." *J. of Ed. Psych.*, 4, 338-352.

Gates, A. I. (22). "The Correlations of Achievement in School Subjects with Intelligence Tests and other Variables." *J. of Ed. Psych.*, 13, Nos. 3, 4 and 5, March, April, May, p. 129 et seq.

Haggerty, M. E. (23). "Intelligence Examination Delta 2." *J. of Ed. Psych.*, 14, 257-277.

Hicks, V. C. (15). "The Value of the Binet Mental Age Tests for First Grade Entrants." *J. of Ed. Psych.*, 6, 157-166.

Holley, C. E. (20). "Mental Tests for School Use." *University of Illinois Bulletin*, Vol. XVII, No. 28.

Hollingworth, L. S. (18). *The Psychology of Special Disability in Spelling*. T. C. Contribs. to Educ. No. 88.

Irwin, E. A. (13). "A Study of the Feebleminded in a West Side School in New York City." *Training School Bulletin*, 10, 65-76.

Irwin, E. A. and Marks, L. A. (24). *Fitting the School to the Child*. New York.

Keener, E. E. (26). "Results of Homogeneous Classification of Junior High School Pupils." *J. of Ed. Res.*, 14, 14-20.

Kempf, G. A. and Collins, S. D. (29). "A Study of the Relation between Mental and Physical Status of Children in two Counties in Illinois." *U. S. Public Health Reports*, 44, 1743-1784.

Lincoln, E. A. (29). "Ability Grouping in Theory and Practice." *Sch. and Soc.*, 30, 447-453.

Maddocks, C. W. (27). "The Factor of Intelligence in School Failures." *Sch. Rev.*, 35, 602-611.

Madsen, I. N. (22). "Intelligence as a Factor in School Progress." *Sch. and Soc.*, 15, 283-288.

Madsen, I. N. (24). "Some Results with the Stanford Revision of the Binet-Simon Tests." *Sch. and Soc.*, 19, 559-562.

Marsden, R. E. (29). "The Intelligence of Children in Some Remote Country Schools." *Forum of Educ.*, 7, 83-90.

McCall, W. A. and Bixler, H. H. (28). *How to Classify Pupils*. Teachers College Bureau of Publications, Columbia University.

Mitchell, D. (22). "Psychological Examination of Pre-School Age Children." *Sch. and Soc.*, 15, 561-568.

Mort, P. R. (28). *The Individual Pupil in the Management of the Class and School*. New York.

Murdoch, K. (22). "The Accomplishment Quotient, Finding and Using It." *T. C. Record.*, 23, 229-239.

O'Flaherty, W. (24). *A Study of the Over-age Child in School Districts 17 & 18*, N. Y. City. Pamphlet.

Paterson, D. G. (18). "A Mental Survey of the School Population of a Kansas Town." *Sch. and Soc.*, 7, 84-89.

Pintner, R. (17). "A Mental Survey of the School Population of a Village." *Sch. and Soc.*, 5, 597-600.

Pintner, R. (18). *The Mental Survey*. New York.

Pintner, R. and Noble, H. (20). "The Classification of Children According to Mental Age." *J. of Ed. Res.*, 2, 713-728.

Pintner, R. and Marshall, H. (21). "Results of the Combined Mental Educational Survey Tests." *J. of Ed. Psych.*, 12, 82-91.

Pressey, S. L. (19). "A Comparison of Two Cities and Their School Systems by Means of a Group Scale of Intelligence." *Ed. Admin. and Super.*, 5, 53-62.

Pressey, S. L. and Thomas, J. B. (19). "A Study of Country Children in (1) A Good and (2) A Poor Farming District, by Means of a Group Scale of Intelligence." *J. of Appl. Psych.*, 3, 283-286.

Purdom, T. (29). *The Value of Homogeneous Grouping.* Baltimore.

Rogers, A. L. (28). *Educational Guidance of Pupils in Private Schools.* Pamphlet No. 3. Private School Teachers Assoc., Philadelphia.

Snoddy, G. S. and Hyde, G. E. (21). "Mental Survey of Utah Schools." *Bulletin of the University of Utah,* Vol. 12, No. 6.

Stalnaker, E. M. and Roller, R. D. (27). "A Study of One Hundred Non-Promoted Children." *J. of Ed. Res.*, 16, 265-270.

Stebbins, R. and Pechstein, L. A. (22). "Quotients I. E. and A." *J. of Ed. Psych.*, 13, 385-398.

Terman, L. M. (17). *The Stanford Revision and Extension of the Binet-Simon Scale.* Baltimore.

Terman, L. M. (19). *The Intelligence of School Children.* Boston.

Terman, L. M. (20). "The Use of Intelligence Tests in the Grading of School Children." *J. of Ed. Res.*, 1, 20-32.

Terman, L. M. et al. (22). *Intelligence Tests and School Re-organization.* Yonkers.

Theisen, W. W. (22). "Relative Progress of VII B Groups Selected on the Basis of Ability." *J. of Ed. Res.*, 5, 295-305.

Thomson, G. H. (21). "The Northumberland Mental Tests." *Brit. J. of Psych.*, 12, 202-222.

Torgerson, T. L. (22). "The Efficiency Quotient as a Measure of Achievement. *J. of Ed. Res.*, 6, 25-32.

Torgerson, T. L. (26). "Is Classification by Mental Ages and Intelligence Quotients Worth While?" *J. of Ed. Res.*, 13, 171.

Van Wagenen, M. J. (26). *Educational Diagnosis and the Measurement of School Children.* New York.

Van Wagenen, M. J. (29). "Comparative Pupil Achievement in Rural, Town and City Schools." University of Minnesota Press, Minneapolis.

Whitcomb, M. E. (22). "Intelligence Tests in the Primary Grades." *J. of Ed. Res.*, 5, 58-61.

Wood, E. P. (29). "The Educational Achievement and Intelligence of Independent School Children." *Ed. Records Bull.*, No. 2, N. Y. Educ. Records Bureau.

Woody, C. (24). "Results Obtained from Ability Grouping." Bureau of Ed. Reference and Research. *Univ. of Mich. Bull.*, No. 72.

Woody, C. (28). "Results of the Testing Program in the Grosse Ile Public School." Bureau of Ref. and Res. *Univ. of Mich. Bull.*, No. 115.

Worlton, J. T. (27-28). "The Effect of Homogeneous Classification on the Scholastic Achievement of Bright Pupils." *Elem. Sch. J.*, 28, 336-345.

CHAPTER XI

THE HIGH SCHOOL PUPIL

Selection in the High School.—We say that the high school is selective, meaning thereby that there are factors at work which select out the pupils who go to high school. This is in contrast to the unselective nature of the elementary school, which is attended by practically all children. In the past the American high school was a much more stringent selective agency than it is at present, because the proportion of children attending high school has increased very greatly. Thorndike (23) estimates that in 1890 only 1 in 10 attended high school, whereas in 1918 this proportion had increased to 1 in 3. The percentage of the population of high school age enrolled in school has increased as follows:

Year	*Per Cent Enrolled*
1910	12.66
1915	17.76
1920	28.43
1926	46.73

Our interest in the selective nature of the high school lies in the fact that there is evidence that this selection was, and still is, partly on the basis of intelligence. On the average the more intelligent tend to go on to high school and the less intelligent to drop out before high school is reached. Dearborn (28) shows how the median I.Q. in three average New England towns rises from 97 in Grade II to 104 in Grade VIII and then continues from 106 in Grade IX to 119 in Grade XII. In St. Louis (Report, 25) the median I.Q. of 8,998 Grade I

children is 100, and the median I.Q. of 6,567 high school freshmen is 103. Again, comparisons of the scores of high school pupils on Army Alpha with the scores made by the men in the army show (Thorndike, 23) that 97 per cent are above the army median. Strickland (26) finds that 99.7 per cent of 1,021 high school seniors fall in the upper half of the army distribution.

There is no doubt therefore that the high school selects from the point of view of intelligence. As the proportion of children going to high school increases the average intelligence diminishes.

General Surveys.—One of the most extensive surveys of high school pupils by means of intelligence tests is that undertaken by Book (22), who tested 6,188 seniors in 320 high schools in the State of Indiana. This cross-section of the high school seniors of the state shows some interesting results. For purposes of rating, arbitrary intelligence grades A, B, C, D, E, F are assigned and the percentage of seniors at each grade is as follows:

Intelligence Grade	A	B	C	D	E	F
Percentage	8	14	52	13	12	1

As in the elementary schools, so also in the high schools we find enormous differences in the mental make-up of different schools and of different communities. To quote the author:

"1. In some schools all members of the senior class possess a superior or very high grade of intelligence. That is to say, all members of the class will be rated A or B, or A, B and C +.

"2. In other schools all members of the senior class rank very low, none scoring above the median for the state. In a few schools the entire senior class would make scores which entitled them to only a D, E or F intelligence rating.

"3. In still other schools a large proportion of all the senior class may possess a C, or average grade of mental ability.

"4. In a fourth type of school there is a marked irregu-

larity among the members of the senior class. Some individuals possess very superior mental ability. Other members of the same class will merit a ranking of E — or F. This situation is much more likely to be found in the smaller high schools. The larger high schools are better graded and seem to have eliminated all inferior students before they reach the senior year."

Comparing communities instead of individual schools, great differences in intelligence are likewise found. Book gathered information from the pupils as to whether they were planning to go to college or not, and he found that pupils of all grades of mental ability from the highest to the lowest were planning to continue their education in college in about equal numbers. "Almost as many students possessing E and F grades of intelligence are going to college as merit a ranking of A + or A." Unfortunately for the colleges many of the brightest students were not planning to go to college at all. Of those rated A +, 22 per cent, A 24 per cent and B 28 per cent did not expect to go to college; while of those ranking D and E, 64 and 62 per cent respectively intended to go to college. The average rating of those intending to go to college is only slightly higher than that of those not intending to go. We seem to be failing, therefore, in seeing to it that the best intelligence of the high school is given a college education. This is surely important, regardless of what our attitude may be with reference to encouraging or discouraging those of inferior intelligence in their college intention.

The school progress of these high school students reveals the fact that in general promotion is controlled by the calendar rather than by ability or accomplishment, and this to a greater extent than in the elementary schools. "It seems to be a habit of high school officials to keep their students in the high school for four years regardless of their ability to do the work." And again, "Most individuals possessing superior or very superior intelligence have been only regularly promoted by the high school."

An analysis of the vocational ambitions of these seniors is interesting and shows the great need of advice and guidance. As far as their choices are concerned, we note that the average score of the boys selecting science, the ministry and journalism is the highest; whereas those selecting medicine, business, farming, stenography or a skilled trade rank lowest. In this connection it is interesting to note the low position of medicine in view of the relatively low standing of the medical officers in the army. There are many other items of interest in Book's survey and we shall have occasion to refer to some of them later on.

Another state-wide survey of high school seniors has been made in Massachusetts by Colvin and MacPhail (24). Intelligence tests were given to 3,333 seniors in 34 schools chosen as representative of all the high schools of the state. The authors conclude that 27 per cent of the seniors are good college risks, 21 per cent are doubtful, and the remaining 52 per cent are bad risks. Unlike Book, these authors found that those planning to continue their education after graduation from high school rated somewhat higher in intelligence; 34 per cent appeared to be good college risks. On the other hand, 16 per cent of those making good scores were not planning to continue their education. As in all surveys, so in this one, we find very great differences between different schools, and also differences between different sections of the state.

In Ruch's (23) survey in Iowa 1,527 high school seniors were given Part I of the Thorndike Intelligence Test. Thirty representative high schools were chosen, including about one-fourth of all Iowa high school seniors. The mean score for each school is given, and the difference between the school ranking highest and the one ranking lowest is 1.5 times the standard deviation of the total distribution of the pupils. The means for the schools range from 101 to 74, and the individual scores from 42 to 145. In addition to the intelligence tests, educational tests were also given. These tests showed less variability than the intelligence tests and the author argues that "the net re-

sult of the educational system is to decrease the amount of individual differences in the raw products [the pupils] on which the schools work." The educational system seems to level individuals toward mediocrity.

Differences in the intelligence of different cities and in different schools of the same city are shown by Madsen's (21) survey by means of the Army Alpha test of 7,168 students in four cities. The median scores reported by him are as follows:

				Omaha		
	Madison	*Rockford*	*Sioux City*	*Central*	*Commerce*	*South*
Seniors	121	124	138	135	111	110
Juniors	122	120	130	132	102	118
Sophomores	109	107	122	122	94	114
Freshmen	96	94	108	107	85	93

In each city there is an increase in score as we ascend from the freshmen to the seniors, with the exception of the Madison seniors. The less intelligent students are eliminated from year to year. The differences between cities are marked, Sioux City being much superior to Madison and Rockford, and showing about the same median scores as the Central High School at Omaha. The differences between the three Omaha schools are greater than those between the first three cities.

Keener's (25) results for one city emphasize the great differences between individual pupils and schools. He reports the scores of 12,652 high school freshmen in Chicago tested by the Otis S. A. Test. The median C.A. is 14 – 5 and the median M.A. is 14 – 4, while the range of M.A.s is from 7 – 6 up and beyond 18 – 6. The median M.A. for schools ranges from 12 – 3 to 15 – 5. The 75th percentile of the lowest school is below the 25th percentile of the highest school. He also shows the well-known negative relationship between C.A. and M.A.; the younger pupils being in general the brighter and the older the duller.

Comparison of City and Country Children.—A comparison of city and country high school children shows a similar

superiority of the city children as we found in the case of elementary school children. Hinds (22) gives the following results in terms of the coefficient of brightness on the Otis test:

	Median Intelligence	*No. of Schools*
City High Schools	100.5	164
Affiliated Town Schools	98.0	290
Small Town Schools	84.4	59
Rural Schools	77.0	68

The author says, "These tests represented as nearly average groups of students in each type of school as it was possible to obtain. This being the case, the conclusion seems to be justified that the country child is lower in general mentality, as measured by the group mental test, than the city child. . . . It should not be forgotten, however, that for years there has been a steady migration of the country's best to the city. Lincolns come from the rural districts, but they never go back."

Book's (22) survey of high school seniors also shows that rural high schools rate decidedly lower than the city high schools and this is true in every section of the state. The median scores on his test are as follows:

	Northern	*Central*	*Southern*
City	141	141	136
Rural	134	135	130

He warns us, however, that we must not forget that children of all degrees of intelligence are found in all types of schools.

Colvin and MacPhail (24) also find a lower median score for their seniors in rural schools. They divide their schools as follows:

	Median Score	*Schools*	*Pupils*
Cities	45.7	17	2,880
Towns of 5,000 up	45.6	6	288
Towns under 5,000	41.9	11	165

They raise the question as to whether the low score of the rural pupils is due entirely to lower intelligence or to a lack of educational advantages enjoyed by city pupils, but they do not discuss the problem.

This same superiority of urban children is also seen in the recent results of Toops who reports on 29,375 Ohio high school seniors as follows:

District	*Mean Intelligence Score*	*n*
County	114.5	12,340
City	122.8	12,690
Private and Parochial	139.4	696
Cleveland	139.3	3,649

Whatever the cause, then, we seem to find a slight inferiority in intelligence among rural children, and this is true both in the high school and the elementary school. This same factor occurs in a later chapter where intelligence is distributed according to parental occupation, and we shall note there the relatively low rating of farming as contrasted with other occupations. The usual explanation of this small difference in intelligence is that there exists a tendency for superior intelligence to migrate from the country to the city. No definite proof of this exists, but it seems a plausible explanation.

Private Schools.—As with the reports of elementary school pupils in private schools, so also do the reports of high school pupils in private schools show very high intelligence. The Army Alpha was used by Anderson (23) with 278 boys in a Connecticut private school and also by Jones (26) in a New York private school. The median scores by grade in these two schools are very high and are about the same as those found among college students. The percentages of A letter grades on the Army Alpha compared with high schools and colleges are:

	Private High School	*Public High School*	*College*	
Grade XII	83	29	60	College Seniors
Grade XI	83	25	58	College Juniors
Grade X	70	17	52	College Sophomores
Grade IX	34	8	45	College Freshmen

Similarly Hildreth (28) at Lincoln School and Flemming (25) at Horace Mann School, private schools connected with Columbia University, show the very high intelligence of the high school group. Hildreth with the Thorndike Intelligence Examination finds 6 per cent poor college risks, 17 per cent fair, 50 per cent good and 26 per cent superior. Flemming uses several different tests and finds on the Terman Group Test that about 90 per cent of the private school children exceed the grade median for ordinary high school pupils. In terms of I.Q.s on this test 67.5 per cent are above I.Q. 110. Wood (29) gives the following median I.Q.s for children tested in various private schools in the East:

	Median I.Q.	*Per Cent Above 110 I.Q.*	*n*
Grade IX	117	78.8	180
Grade X	117	75.7	177
Grade XI	118	82.0	150
Grade XII	117	75.0	136

All of these reports and many others show that private schools are attracting a very superior type of intellect. They are dealing essentially with superior children and their educational problems are, therefore, somewhat different from those confronting the public high school.

Continuation Schools.—Pupils in continuation schools are about the same age as those in the regular high schools. Three reports show median I.Q.s for various groups of continuation school pupils:

Author	*I.Q.*	*n*
Hopkins (24)	86	1,197
Sudweeks (27)	86	1,824
Martin (28)	93	3,125

The work of Hopkins (24) covered all the 14- and 15-year-olds in several communities, and also a large sampling of children of like age in the regular school. He gives this comparison of median I.Q.s for continuation and regular school children:

	Continuation	*Regular*	*Combined Group*
Age 14	85	102	98
Age 15	87	104	97

Since these results are based upon practically all the 14- and 15-year-olds in certain communities, we see clearly how the regular school is selecting on the basis of intelligence and leaving the weaker in intelligence to the continuation school. Sudweeks (27) gives a percentage distribution of I.Q.s:

I.Q.	*Percentage*
120-140	0.05
110-120	1.2
90-110	32.0
80-90	39.4
70-80	22.2
70 and below	5.1

Here we note how the great mass of pupils lies below I.Q. 90. Note also that only 1.2 per cent have I.Q.s above 110 and compare this with the 70 to 80 per cent of I.Q.s above 110 in the private schools. Martin's (28) data are all from one city and seem to indicate a slightly better type of intelligence in the continuation schools of that city. The general conclusion from all these results is that the continuation school contains the weaker intellect that is not able to survive in the regular school.

Courses of Study.—Intelligence varies according to the course of study elected in high school, those scoring highest tending to be found in the more abstract courses and those scoring lowest in the more practical courses, with much overlapping among all courses. Clark (24) gives the average I.Q.s of those graduating from various courses:

Course	*Average I.Q.*	*n*
General	114.5	151
Commercial	109.4	102
Technical	108.9	100
Industrial Art	103.1	34
Dressmaking	97.4	51

The St. Louis Report (25) gives the median I.Q.s for those entering various courses:

Course	*Median I.Q.*	*n*
Scientific	109.8	228
General	106.3	2,107
Classical	106.1	52
Commercial	103.2	806
Manual Training	102.5	655
Art	102.1	235
Home Economics	100.5	296

Two year courses in practical subjects follow with still lower I.Q.s.

Keener (25) gives the median mental ages of pupils selecting various courses:

Course	*Median M.A.*	*n*
Latin	15-1	625
Modern Language	14-11	403
Commercial	14-9	311
Technical	14-5	1,394

Two year courses follow with lower M.A.s.

Colvin and MacPhail (24) give the median scores:

Course	*Median Score*	*n*
Classical	56.5	54
Academic	55.0	34
College Preparatory	53.6	682
Scientific	52.9	225
Normal Preparatory	43.9	150
General	43.5	421
Vocational	43.3	188
Commercial	40.9	1,261

The general picture resulting from these various studies is that of a decrease in intelligence as the course becomes less abstract or more practical in nature. The differences in intelligence ratings between courses are not large, but the consistency of these differences in the various reports is marked. Some of this difference is probably due to the verbal type of intelligence test generally used, favoring those electing the more verbal type of course; but all of the difference is not likely due to this. The more intelligent are likely to be attracted to and survive in the traditionally "harder" courses.

School Marks.—The general intelligence of a pupil should determine to some extent the type of work he does in school. We have seen how general ability controls the pupil's coming to high school in the sense that those of low ability are eliminated, and now among the select group remaining we shall see that it still plays a part in determing his standing in accordance with the marks given to him by his teachers. Many correlations between intelligence test scores and total high school marks have been computed. Table 8 shows a sampling of the correlations reported by various workers. Most of these lie between .4 and .5 and this would seem to represent the ordinary agreement between the usual group intelligence test and some composite of school marks. This correlation is not high for purposes of prediction but it shows the substantial agreement which exists between general ability as measured by our

present tests and the rather unreliable school marks as allotted by various teachers.

TABLE 8

Correlations Between Intelligence Scores and School Marks

Author and Date	*r*	*Remarks*
Madsen (21)	.40	428 pupils in Grades IX to XII
West (21)	.60	173 pupils in Grades IX to XII
Proctor (21)	.43	average of 3 r's
Miller (22)	.52	55 pupils in Grade IX
Book (22)	.28	5,748 high school seniors
Book (22)	.47	124 in one school
Jordan (22)	.47	average of 4 r's
Haggerty (23)	.56	Grade IX
Colvin (24)	.43	average of 6 r's
Ross (25)	.37	Grade IX
Flemming (25)	.50	average of 6 r's
Unzicker (25)	.47	Grades VII to IX
Rector (25)	.44	431 pupils
Strickland (25)	.45	935 seniors

Jordan (23) summarized the correlations found by many workers for three tests as follows:

Test	*Median r*	*Range*	*No. of r's*
Army Alpha	.38	.19-.52	26
Otis	.49	.31-.82	16
Terman	.47	.30-.67	9

Special Subjects.—Correlations between intelligence test scores and marks in particular subjects have been frequently reported. Table 9 gives a sample of the results obtained. They are very similar to the results obtained for total school marks. There seems no marked tendency for the correlations to be higher in any one of the usual academic subjects rather than in any other. Thorndike (23) has discussed in detail the relationship of intelligence and algebra, pointing out the fact

TABLE 9

CORRELATIONS BETWEEN INTELLIGENCE TESTS AND SEPARATE SUBJECTS

Subject	*r*	*Author*	*Remarks*
Algebra	.35	Thorndike (23)	Army Alpha
	.50	Crathorne (22)	Two intelligence tests
	.38	Buckingham (21)	Army Alpha
	.46	Proctor (21)	Binet I.Q.s
	.50	Bright (21)	
Mathematics	.37	Book (22)	
	.34	Unzicker (25)	
	.33	Jordan (22)	Average of 4 r's
	.58	Flemming (25)	Terman Test
English	.44	Book (22)	
	.72	Bright (21)	
	.41	Jordan (22)	Average of 4 r's
	.56	Flemming (25)	Terman Test
	.35	Unzicker (25)	
History	.25	Book (22)	
	.28	Jordan (22)	Average of 4 r's
	.47	Unzicker (25)	
	.52	Flemming (25)	Terman Test
Latin	.26	Book (22)	
	.65	Bright (21)	
	.40	Flemming (25)	
Chemistry	.52	Book (22)	
General Science	.47	Jordan (22)	Average of 4 r's
French	.33	Flemming (25)	
Handicraft	.36	Bright (21)	

that in dealing with high school pupils alone, we are dealing with a rather select group of individuals. He says, "We conclude then that the correlation between algebra and intelligence is .80 when taken over the total range of the unselected adult population. For high school freshmen as a group this correlation would be in the neighborhood of .70." After ex-

plaining what this correlation signifies in the sense that this may mean that a few of very high intelligence may be very low in algebraic ability, he concludes, "Let not this emphasis on the extreme cases blind one to the main fact that, by and large, high intelligence means fine ability in algebra and low intelligence means poor ability. Of the twenty-two brightest pupils out of the hundred, seventeen will be above the average in algebra, and of the twenty-two stupidest pupils, seventeen will be below the average." In all probability the situation is much the same in other abstract subjects such as Latin, and slightly less so in the less exact subjects, such as English and History, and much less so in the more practical and commercial subjects.

Practical Use of Tests.—The results of intelligence tests in high school have been used for educational guidance and for classification, and a few such reports may be mentioned here. The most interesting attempt to use the tests for educational guidance is reported by Proctor (21). Children entering the high school indicated the subjects they desired to take, their future educational plans with reference to college, normal school and the like, and also their vocational ambitions. In addition to this, the educational counsellor had a record of their grades in the elementary school as well as their intelligence ratings on the Army Alpha and Stanford-Binet Tests. The tests function as one item in helping to guide the students in arranging their course of study. Proctor gives many interesting samples of individual cases, and then compares the first year's work of a "guided" group of children with another group of "unguided" cases. The facts are as follows in terms of per cent:

	Out at Work	*Out by Transfer*	*Failed 1 Subject*	*Failed 2 or More*
Guided	4.5	9.1	18.2	0.0
Unguided	12.1	13.1	30.8	10.3

The two groups were practically equal in intelligence, having median I.Q.s of 105 and 108. Evidently guidance in the choice

of studies has saved many from failure or elimination from school, and, as the author says, "it is certain that the methods applied in this instance, if employed in any high school, would prove greatly superior to the wasteful trial and error methods that now prevail." In the same study Proctor gives the median I.Q.s for various groups of high school students, as follows:

First year students	105
High School graduates	111
Those going to college	116

The lower limits in terms of I.Q. below which there is little chance for success in high school or college, as at present organized, should be found, in order to help the educational counsellor in his work. Thorndike (23) estimates that in general a pupil with an I.Q. below 110 on the Stanford Revision "will be unable to understand the symbolism, generalizations and proofs of algebra." Similarly Terman (19) says that pupils with I.Q.s below 90 are practically certain to fail in such subjects as algebra and Latin, and are not likely to graduate from high school. Again Cobb (22) says, "Probably in 90 cases out of 100, it is unwise to guide the average or less intelligent than average child into the present academic high school. Unless his I.Q. is over 100, or his M.A. definitely over 14, he should be encouraged to try some other type of training."

The classification of eighth grade pupils on the basis of intelligence tests is reported by Dickson (20) and their success in high school studied. Two groups, "regular" and "accelerated," were formed and the "accelerated" group covered the eighth grade work in one semester. After a year's work in high school the "accelerated" were compared with the regular high school students with whom they had been competing and were found to be quite superior in scholarship, in spite of being 17 months younger. Dickson also reports sections in English, algebra and arithmetic formed on the basis of intelligence tests, and quotes the testimony of teachers with reference to the value of such classification.

There are several other studies describing experiments in ability grouping in which the authors express their opinion as to the advantages of such grouping. Symonds (27) discusses at length the advantages and disadvantages, showing that most of the objections are based on a faulty conception of the educative process.

Conclusions.—Intelligence tests show that the high school pupil belongs in the upper half of the distribution of intelligence. Selection has taken place from the point of view of abstract intelligence. This will be increasingly less true if the high schools continue to include more and more of the children of high school age. We note further that as the high schools have expanded and increased the variety of their courses, the more intelligent tend to be found in the more academic courses and the less intelligent in the more practical courses.

Again we note, as in the elementary school, that high schools vary very much in the average intelligence of their pupils. The highest average intelligence seems to be found in the private schools and the lowest in the rural public high schools.

There is a fair correlation between intelligence scores and school marks, generally between .4 and .5. This will vary greatly according to the accuracy of the intelligence test and the reliability of the marks. Both are defective measures. Such correlations show that intelligence is only one of the important factors making for success in high school studies.

Intelligence tests are useful for general classification purposes and to some extent for educational and vocational guidance. As our high schools are at present, children with I.Q.s between 90 and 100 will have great difficulty in graduating. Those with I.Q.s between 100 and 110 will be able to manage successfully the less abstract courses, and those with I.Q.s above 110 will be successful in the more abstract academic work.

REFERENCES

Anderson, J. E. (23). "The Intelligence of a Highly Selected Group." *Sch. and Soc.*, 16, 723-725.

Book, W. F. (22). *The Intelligence of High School Seniors.* New York.

Bright, I. J. (21). "The Intelligence Examination for High School Freshmen." *J. of Ed. Res.*, 4, 44-55.

Buckingham, B. R. (21). "Mathematical Ability as Related to General Intelligence." *Sch. Sci. and Math.*, 21, 205-215.

Clark, R. S. (24). "Some Results of Psychological Tests Given to Groups of Public School Pupils of New York City." *Contrib. to Ed.*, Vol. I. N. Y. Soc. for the Exper. Study of Ed., 98-116.

Cobb, M. V. (22). "The Limits Set to Educational Achievement by Limited Intelligence." *J. of Ed. Psych.*, 13, 449-464, 546-555.

Colvin, S. S. and MacPhail, A. H. (24). "Intelligence of Seniors in the High Schools of Massachusetts." *Bulletin*, No. 9, Bureau of Education, Washington.

Crathorne, A. R. (22). "A Critical Study of the Correlation Method Applied to Grades"; to be Chap. 10, of *The Reorganization of Mathematics in Secondary Education:* Report of Nat'l Committee on Math. Requirements.

Dearborn, W. F. (28). *Intelligence Tests.* Boston.

Dickson, V. E. (20). "The Use of Group Mental Tests in the Guidance of Eighth Grade and High School Pupils." *J. of Ed. Res.*, 2, 601-610.

Flemming, C. W. (25). *A Detailed Analysis of Achievement in the High School.* T. C. Contrib. to Educ., No. 196.

Haggerty, M. E. (23). "Intelligence Examination Delta 2." *J. of Ed. Psych.*, 14, 257-276.

Hildreth, G. (27). "Thorndike College Entrance Test Results in a Senior High School." *T. C. Record*, 28, 1035-1044.

Hinds, J. H. (22). "A Comparison of the Brightness of Country and City High School Children." *J. of Ed. Res.*, 5, 120-124.

Hopkins, L. T. (24). *The Intelligence of Continuation School Children.* Cambridge, Harvard University Press.

Jones, H. E. (26). "The Intelligence of Preparatory School Students." *J. of Ed. Psych.*, 17, 376-383.

Jordan, A. M. (22). "Correlations of Four Intelligence Tests with Grades." *J. of Ed. Psych.*, 13, 419-429.

Jordan, A. M. (23). "The Validation of Intelligence Tests." *J. of Ed. Psych.*, 14, 348-366, 414-428.

Keener, E. E. (25). "Mental Ability of Freshman High School Pupils." *J. of Ed. Res.*, 11, 113-122.

Madsen, I. N. (21). "Group Intelligence Tests as a Means of Prognosis in High School." *J. of Ed. Res.*, 3, 43-52.

Martin, A. L. (28). *The Child Study Department.* The Work of the Public School. Board of Education, Rochester, N. Y.

Miller, W. S. (22). "The Administrative Use of Intelligence Tests in the High School." *Twenty-first Yearbook N. S. S. E.,* Chap. VII.

Proctor, W. M. (21). "Psychological Tests and Guidance of High School Pupils." *J. of Ed. Res.* Mono., No. 1.

Rector, W. G. (25). "A Study in the Prediction of High School Success." *J. of Ed. Psych.,* 16, 28-37.

Report (25), "Division of Tests and Measurements." *Public School Messenger.* Board of Ed., St. Louis. Vol. 23, No. 1, pp. 45-57.

Ross, C. C. (25). *The Relation between Grade School Record and High School Achievement.* T. C. Contribs. to Educ., No. 166.

Ruch, G. M. (23). "A Mental-Educational Survey of 1550 Iowa High School Seniors." *Univ. of Iowa Studies in Education,* Vol. II, No. 5, University of Iowa, Iowa City.

Ruch, G. M., and Stoddard, G. D. (27). *Tests and Measurements in High School Instruction.* Yonkers.

Strickland, V. L. (26). "Relative Levels of Intelligence in College and High School." *Kansas State Agricultural College Bulletin,* Vol. 10, No. 4.

Sudweeks, I. (27). "Intelligence of the Continuation School Pupils of Wisconsin." *J. of Ed. Psych.,* 18, 601-611.

Symonds, P. M. (27). *Measurement in Secondary Education.* New York.

Terman, L. M. (19). *The Intelligence of School Children.* Boston.

Thorndike, E. L. et al. (23). *The Psychology of Algebra.* New York.

Toops, H. A. (No date). *Ohio High School Bulletin,* No. 2. Mimeographed.

Unzicker, S. P. (25). "Teachers' Marks and Intelligence." *J. of Ed. Res.,* 11, 123-131.

West, R. L. (21). "An Experiment with the Otis Group Intelligence Scale." *J. of Ed. Res.,* 3, 261-268.

Wood, E. P. (29). *The Educational Achievement and Intelligence of Independent School Children.* Educational Records Bureau, New York City.

CHAPTER XII

THE COLLEGE STUDENT

The college student has from the very beginning preeminently been "the dog" upon which the psychologist has experimented in the laboratory. Working almost always in a university the psychologist has found an abundant supply of willing and valuable observers among the student body. The science of psychology is greatly indebted to their help and assistance, and they themselves have likewise profited by this closer and more intimate contact with psychology. It is not surprising, therefore, that we should find mental testing of college students at a very early date in the history of our subject. As we have indicated in Chapter I, the work was begun very early by Cattell, who in 1890 reported results of tests on students at the University of Pennsylvania, and again in 1896 on students at Columbia.

The first tests used were of a decided sensory and motor type and were mostly given as individual tests. Later on more distinctively mental tests were employed, both individual and group, until in recent years the group intelligence test, as we now know it, is most commonly used. Our treatment of the vast amount of work on the testing of the college student must necessarily be somewhat sketchy. It would be impossible and unnecessary to give a minute description of all the studies.

Miscellaneous Group and Individual Tests.—It might be truly said that there is practically no mental test which has not been tried out on college students, and very few upon which we do not have reports of some sort or other. The reports range from a few tests tried out on a few students to many tests tried out on a great number; from studies of an incidental

nature to studies in which a group of students has been intensively tested. All of the studies are more or less concerned with the reliability of the tests, but particularly with the problem of the value of psychological tests for predicting success in college work. We find, therefore, a great many correlations between test results and academic grades. A sampling of such correlations is given below. In this list the coefficient of correlation is often an average of several coefficients between tests and different academic subjects. Sometimes it refers to only a few subjects and a limited class of students, such as a group of freshmen; at other times it may be obtained from a combination of many academic grades or may contain students of several college years. All of these factors will, of course, affect the coefficient.

Correlation Between Tests and Academic Grades

Author	*Coefficient*
Wissler (01)	.09
Calfee (13)	.23
Rowland and Lowden (16)	.37
Waugh (16)	.41
Bell (16)	.31
Hollingworth (16)	.50
Kitson (17)	.44
King (17)	.27
Haggerty (18)	.65
King and McCrory (18)	.38
Uhl (19)	.43
Rogers (20)	.55
Carothers (21)	.21

It will be noticed that the great majority of coefficients are below .50. Coefficients of correlation between certain tests or particular groupings of tests and specific academic subjects vary around the coefficients given in this list from negative coefficients up to high positive coefficients of about .90, but in

general the relationship between psychological tests and academic success is not very high. There are many other important elements entering into success in college studies besides the one of intelligence. Furthermore, the college student is very homogeneous in general intelligence as compared with the population at large, and this homogeneity of the group reduces the correlation. Nevertheless, as many of the authors point out, the results of judiciously chosen tests may help very decidedly, along with other facts about the student, to advise the student as to his work and to help the administrative officers in problems of dismissal or suspension. Kitson's (17) suggestions in reference to vocational guidance are valuable here. Carothers' (21) psychographic charts show well the variations in different abilities in the same individual, and her suggestions as to personal interviews with students to discuss the results of the tests are valuable.

Some work has been done in comparing the results of tests with estimates of intelligence. Usually these estimates are the ratings of students by one or more instructors. Some of the results obtained are:

Correlation Between Tests and Estimated Intelligence

Author	*Coefficient*
Hollingworth (16)	.61
Kitson (17)	.57
Uhl (19)	.36
Thurstone (19)	.60

These coefficients are of much the same magnitude as those between tests and academic grades. Estimates of intelligence, as well as academic grades, are themselves very unreliable and this unreliability tends to lower the correlations. Thurstone (21) says, "I have found that instructors' estimates, on account of their unreliability, are, in general, unsuitable as a criterion by which to judge the predictive value of a mental test. The several instructors' estimates for the same student vary con-

siderably more than the corresponding scholarship grades."

Two studies using the Binet Tests with college students are reported. The number of Binet tests valuable for adults is small. Their wide dissemination among college students makes them of little general value. They do not seem to correlate any higher than the general run of tests. Caldwell (19) reports a correlation of .44 between the Stanford-Binet tests and academic grades. Between the same tests and estimates of intelligence, a correlation of .47 is reported by Caldwell (19), and of .53 by Downey (17). Caldwell finds the I.Q. procedure for adults unsatisfactory and says "there should be some means for denoting higher adult I.Q's." But, surely, it would seem wiser to abandon the I.Q. procedure altogether for adults and particularly for college students.

The Army Alpha Test.—During the war, and for several years thereafter, Army Alpha was used widely in colleges and universities. Table 10 gives median scores reported by many schools.

This list does not pretend to be exhaustive. Undoubtedly many more groups of college students have been tested and many workers have not published their results. The list, however, gives a fair indication of the results in general. There is naturally some difference in the medians of the various groups. Some of this difference may be due to the composition of the groups, whether made up of freshmen, seniors and the like, or otherwise selected; some of it may be due to a difference in the form of the test used; but some of it is undoubtedly due to the different caliber of the student body found in different colleges. At the bottom of the list is given the median score for officers in the army, showing that the mental rating of officers resembles the rating of college students, as we should expect. We notice that the college medians in Table 10 range from 80 to 160. The central tendency is somewhere between 130 and 140, and this is the same as the median score for the officers. The average score for the drafted man is about 60. Out of

twenty-nine thousand men only about one per cent reached or exceeded a score of 135. It is obvious, therefore, that college students are a highly selected group in regard to the type of intelligence measured by Army Alpha.

TABLE 10

ARMY ALPHA: MEDIAN SCORES IN VARIOUS COLLEGES

College	*Median*	*No. of Cases*
Yale (men)	160	406
Massachusetts Agricultural (men)	150	154
Oberlin (both sexes)	149	330
Brown University (men)	142	210
Syracuse University (both sexes)	142	786
Colorado College (men)	142	148
Colorado College (women)	142	178
Dickinson (both sexes)	141	72
Rutgers College (men)	138	358
Johns Hopkins (men)	137	140
Notre Dame (men)	137	321
Ohio State University (both sexes)	136	5,950
Newcomb (women)	135	558
Lewiston Normal (both sexes)	133	693
Penn State College (both sexes)	132	847
University of Minnesota (men)	129	534
University of North Dakota (women)	129	117
University of Minnesota (women)	128	354
Southern Methodist (men)	127	162
University of Idaho (men)	125	277
Purdue (men)	124	1,162
Southern Methodist (women)	123	159
Colorado Teachers College (women)	122	266
University of Florida (men)	120	215
University of Idaho (women)	117	169
Lincoln Memorial (men)	86	171
Atlanta Southern Dental (men)	80	184
White Officers in the Army	139	15,385

A comparison of the median scores of students in different colleges in a university is given by Noble and Arps (20), and we may compare their results with those reported for a few colleges at the University of Illinois:

	Ohio State		*Illinois*	
	Median	*Cases*	*Median*	*Cases*
Graduates	157	152	154	161
Commerce	147	52	143	539
Medicine	142	141	...	...
Law	142	141	...	...
Engineering	141	1,392	144	755
Agriculture	133	859	139	385
Arts	133	1,966	145	1,410
Education	133	382	...	...
Pharmacy	125	109	...	...
Dentistry	115	152	...	...
Vet. Medicine	112	93	...	...

Both reports place the graduate schools at the top of the lists by a very appreciable margin. There is no agreement between the two reports as to the standing of the four other colleges, where we have comparable data. At Ohio State the professional schools range themselves about equally above and below the median for the Arts College. Undoubtedly there are selective influences at work in determining the student body of a college and these influences may not be the same at each university. At Ohio State the relatively low standing of the students in Dentistry and Veterinary Medicine corresponds with the low standing of dental and veterinary officers in the army. The medical group, however, is above the engineering group at the university, which was notably not the case in the army.

How do the intelligence ratings on the army tests correlate with the students' work in college as evidenced by his class marks? Numerous correlations are reported:

	Coefficient	*No. of Cases*	*Class*
Bridges (20)	.35	436	All classes
	.15	36	Graduates
	.38	100	Seniors
	.35	100	Seniors
	.29	100	Juniors
	.29	100	Sophomores
Colvin (19)	.45	212	Freshmen
Stone (22)	.44	633	Freshmen
	.33	633	Freshmen
	.50	622	Freshmen
Van Wagenen (20)	.46	84	Sophomores
	.50	84	Sophomores
DeCamp (21)	.41	320	Freshmen
Burtt (22)	.38	651	Seniors
Miner (27)	.29	37	Engineers
Viteles (22)	.21	59	Business

These are typical of the results usually obtained in comparing the intelligence ratings with academic grades for all subjects for one or more semesters. The lowest correlation is .15, but this seems to be a decided exception. In general they fluctuate between .30 and .50. If we compare these coefficients with the coefficients for the miscellaneous groups of tests given on page 293, we note a much greater uniformity in the results from the army tests. In both lists, however, most of the coefficients lie between .30 and .50. The army tests do not seem to correlate more highly than some other groups of tests. They have, however, the advantage over most of the tests previously reported in being easy to give and score and economical in time.

Stone (22) reports coefficients of correlation between the army test and separate freshman subjects. They are all positive and range from .11 to .50. He also attempts an analysis of the different tests of the whole scale, and concludes, "all in all, the present Alpha would, from the standpoint of elective ad-

visory purposes, seem to be as random an agent as the traditional campus method of selecting courses." Davis (21), using the regression equation obtained after correlating the army test with total grades, would make use of it in estimating whether a man is working up to his ability or not, but reports no practical use of this scheme. The most pessimistic attitude with reference to the use of the army tests is expressed by Bridges (22), who summarizes his experience with the work at Ohio State and concludes that the tests at present are of practically no value in university work. Instead of these general group tests, he recommends detailed psychological and psychiatrical examinations of special problem cases. Van Wagenen (20) is much more hopeful as to their possible value in college work. His analysis brings out the fact that "over fifty per cent of the students will not change their standing in academic marks from their standing in the army test by more than one-half sigma unit or by more than one-fifth of the range of the scores of college students or by more than a change from a D to a C, or from a C to a B, as these marks are given to college students." He feels, however, that the tests "prove least useful just where reliable results from their use are most needed; namely, in eliminating those most likely to fail in their college work and in selecting for special groups those who are most likely to attain the higher degrees of success." And further, "there is yet, however, far too large a discrepancy between the positions attained in the Army Tests and those achieved in academic marks to warrant the use of the Army Tests for purposes of rigid selection." Van Wagenen is hopeful that modifications and changes in the tests may lead to greater predictive value.

This work with the army tests in colleges and universities, although extensive, has been more or less experimental in nature. Unpublished accounts of the work in several universities indicate that practical use is being made of the results to some extent. Men with high intelligence ratings and low college grades are reported to improve frequently when confronted

with the facts. Deans and other administrative officers take the intelligence rating into consideration when dealing with academic delinquents and cases of probation, dismissal, reinstatement, petitions to carry extra work and the like. The extent to which students of low intelligence are consuming administrative time and energy is shown by the report from Ohio State, in which 81 per cent of the delinquents in a certain college were found to be students with intelligence scores equal to those received by the lowest 5 per cent of the whole student body.

The Thorndike Tests.—The Thorndike Intelligence Examination is very much more comprehensive and very much more difficult than the Army Alpha. It requires two hours and fifty minutes actual working time and includes some educational tests of a type suitable for high school graduates. It is used as a part of the entrance requirements at Columbia College. A correlation of .65 between the work of the entire freshman year and the scores on the mental test is reported by Thorndike (21). The test is obviously detecting the poorer students, for "of eleven boys at Columbia, reported to the dean's office for inability to do college work in the early weeks of the year, all had notably low scores in the intelligence examination. Of a score or more so reported as a result of the midterm records, all but two had low scores. The defective college work of these two was by common consent not due to intellectual defect" (Thorndike, 20).

Not only for admission, but also for administrative purposes are the tests proving valuable. Thorndike (21) quotes the Dean of the College as follows:

"In addition to the use of the results of the mental tests in admission to college, they have been most helpful in my office as an aid in arriving at a diagnosis of academic maladies. A boy who has a poor academic record and a low mental test grade generally needs very different treatment from the student whose record is poor but whose mental test mark is high. And in several cases the mental test has afforded the clue which

has enabled my office in cooperation with the university physician so to advise the boy that he has not only escaped being dropped, but has become an excellent academic citizen."

Wood (23), in describing the work at Columbia, says, "In 1919 the office of Admissions at Columbia University inaugurated an experiment for the purpose of discovering the value of the Thorndike College Entrance Intelligence Examination for High School Graduates as a criterion for admission to Columbia College. From an attitude of healthy if not severe skepticism toward the use of intelligence tests for this purpose, the whole college administration came, within the space of two years, to consider the intelligence tests as an indispensable part, not only of the admission machinery, but also of the administration of the college in the Dean's Office." Wood also gives the correlations between the intelligence scores and scholarship records of 111 students who remained in college during two years:

	r
Freshman Year	.63
Sophomore Year	.62
Freshman and Sophomore Years	.67

That the intelligence test has higher predictive value of a man's work in college is shown by the following correlations with the scholarship record for two years:

	r	*n*
Thorndike Intelligence Test	.67	111
Regents Examinations	.64	144
Secondary School Marks	.26	103

College Marks.—Innumerable coefficients of correlation have been published between intelligence scores and marks in academic work. MacPhail (24) has given a summary of these results up to 1924, and this is reproduced in Table 11. The reader interested in the bibliographic references for all these findings should consult MacPhail (24). As MacPhail points out, the central tendency of all these correlations lies between .40 and .45. There are few below .30 and few above .60. Over

TABLE II

CORRELATIONS BETWEEN INTELLIGENCE AND ACADEMIC MARKS *

Institution	*Test Used*	*Coefficient of Correlation*
Arkansas University	Alpha	.28-.65
Barnard College	Local	.14-.27
Beloit College	Various	.24-.54
Boston University College.	Boston University	.64
Brown University	Alpha	.44-.46
Brown University	Thorndike	.37-.53
Brown University	Brown University	.34-.60
Brown University	Brown and Thorndike combined	.36-.46
Carnegie Institute	Alpha	.34
Carnegie Institute	Thurstone IV	.37
Carnegie Institute	Combination of 5 tests	.60
California University	Thorndike	.46
Chicago University	Kitson	.20-.44
Chicago University	Thorndike	.41
Columbia University	Thorndike	.59-.67
Cornell	Thurstone	.31
Cornell	Terman Group	.44
Dartmouth	Alpha	.44
Dartmouth	Definition	.55
Evanston	Thurstone	.29
Goucher College	Thorndike	.46-.67
Goucher College	Thurstone	.20-.26
Hamline University	Alpha	.47
Hood College	Terman	.48
Hood College	Rogers	.34
Hood College	Average of Otis, Terman, Thurstone, Rogers	.45
Illinois State Normal University	Alpha	.48

* From MacPhail, A. H. The Intelligence of College Students. By permission of Warwick and York, publishers.

Institution	*Test Used*	*Coefficient of Correlation*
Illinois University	Alpha	.37
Illinois University	Thurstone IV	.48
Illinois University	Illinois University	.57
Minnesota University (Medical School)	Reading and Omnibus	.65
Minnesota University	Alpha	.13-.50
Montana State University.	Smith	.36
Montana State University.	Thorndike	.46
Ohio State	Thurstone IV	.41
Ohio State	Alpha	.15-.38
Oregon University	Alpha	.49
Pennsylvania State	Alpha	.41
Pennsylvania State	Thurstone IV	.32
Pennsylvania State	Binet-Simon (SR)	.17
Rutgers College	Thorndike	.45
Randolph-Macon	Stanford-Binet	.44
Reed College		.37
Rochester University	Alpha	.30
Simmons	Roback	.50
Smith College	Rogers	.40
Southern Methodist University	Alpha	.52
Stanford University	Terman Group	.42-.54
Stanford University	Stanford-Binet	.42-.46
Stanford University	Alpha	.31-.43
Syracuse University	Alpha	.20-.40
Texas University	Special Test	.32
Vassar	Thurstone	.33
Washington University ..		.52-.71
Wyoming University	Stanford-Binet	.53
Wyoming University	Thorndike	.40
Yale University	Local	.50
Yale University	Alpha	.37-.38
43 institutions	Thurstone IV	.29

TABLE 12

ADDITIONAL CORRELATIONS BETWEEN INTELLIGENCE AND COLLEGE MARKS

Author and Date	*r*	*n*	*Test*	*Students*	*College*
Root (23)	.51	569	Thorndike	Freshmen	Pittsburgh
Perrin (24)	.49	792	Special	Freshmen	Texas
Johnson (24) ...	.50	435	Special	Freshmen	Minnesota
Symonds (24) ...	.41	212	Thorndike	Freshmen	Hawaii
MacPhail (24) ..	.37 to .46	various	Brown University	All years	Brown
Rogers (24-25) ..	.61	254	Thorndike	Freshmen	Goucher
Anderson (26) ..	.30 to .41	various	Special	All years	Yale
Davidson (27-28)	.47 to .57	various	Brown University	All years	Brown
Guiler (27)	.40 to .52	80	Various	All years	Ohio
Kornhauser (27) .	.42 to .51	104 to 143	Various	Commerce	Chicago
Crane (27)	.27 to .43	various	Thorndike or Thurstone	All years	Bryn Mawr
Cleeton (27)	.38 to .52	95 to 257	Thorndike	First two years	Carnegie Tech.
Toll (28)	.19 to .42	various	Various	All years	Amherst
Condit (29)	.45	559	Thurstone	Freshmen	Colorado
Brigham (29) ...	.51 to .55	300	C.E.B. Aptitude	All years	Not given
Crawford (30) ..	.46 to .49	147 to 606	C.E.B. Aptitude	Freshmen	Yale

two thirds lie between .30 and .50. In Toops' (26) summary of the status of intelligence testing in colleges, he reports a median correlation of .46 based on reports from 43 colleges, and this confirms the findings of MacPhail. Table 12 gives additional correlations since MacPhail's study.

All these correlations between marks and intelligence scores show conclusively that intelligence is one of the most important factors making for high marks. It is, perhaps, the one most important characteristic of the individual so far as marks are concerned. At the same time, it is clearly demonstrated that it is by no means the only factor. Other factors, such as industry, zeal, interest, health, and so forth, also influence the marks a student will obtain.

Courses and Subjects.—Several reports have dealt with the intelligence of students in various courses, but there is very little agreement among them. Perhaps there is a slight tendency for liberal arts students to score higher. Again within a given college there may be a difference in the intelligence of students electing various major studies. Thus Rogers (24) finds that Natural Science students have the highest scores, followed by English, Foreign Languages and Social Science. On the other hand Dexter (29) finds the situation almost reversed, for he ranks the majors as follows: English, Social Sciences, Foreign Languages, History, Natural Sciences. Evidently we can make no generalizations about the relative intelligence of students in various colleges or in various major courses. This varies widely from college to college and from course to course.

Several reports give correlations between intelligence scores and marks in a particular subject. Thus Perrin (24) finds the following correlations:

Subject	*r*	*n*
Mathematics	.80	482
Romance Languages	.79	612
Physics	.78	104
Zoology	.78	147

Subject	*r*	*n*
History	.76	425
English	.72	780
Botany	.72	115
Chemistry	.69	271
Education	.66	274
Geology	.65	105

Root (23) using the Thorndike Tests finds the following correlations for various classes:

Subject	*r*	*n*
Biology	.50 to .53	73 to 151
Chemistry	.43	258
English	.36	495
French	.40 to .45	72 to 131
German	.50	60
History	.43 to .48	48 to 122
Human Progress	.69	527
Mathematics	.39 to .61	23 to 110
Physics	.50	134
Spanish	.47 to .67	14 to 71

Correlations of intelligence scores with marks in psychology classes have been reported:

Spence (27-28)	.61 and .42
Nelson (27)	.51, .64 and .77
Miller (25)	.37

These must suffice as samples of this type of work. Evidently it is impossible to say that any one subject rather than any other correlates higher with intelligence. Classes in the same subject in the same college may vary greatly in their correlations. The spread of intelligence within any class, the reliability of any individual instructor's marks and many other factors will affect the correlation. Where we have very reliable measures of both intelligence and class achievement as in the Spence report, we find the correlation as high as those usually reported for the total grades for a semester.

Age and Intelligence.—As in the lower schools, so also in the college, we find a slight negative correlation with age. The younger student is likely to be the more intelligent. Brigham (27) finds a negative correlation of .28 on the C. E. B. Scholastic Aptitude Test. Terman (23) shows how the Thorndike scores tend to decrease with age. This is more marked for the men than for the women at Stanford. His results for first year men are:

Age	*Ave. Score*	*n*
15	76	9
16	78	51
17	76	227
18	72	355
19	67	198
20	69	85
21	66	50
22 +	63	109

And Whinery (26) compares the age of the top and bottom five per cents on the intelligence test of Ohio State freshmen. He finds the average age of the top five per cent to be 18.16 and the bottom five per cent to be 19.52, showing a difference of 1.36 years.

Fraternity Men.—In general the results seem to show slightly lower intelligence scores for fraternity as contrasted with non-fraternity men. MacPhail (24) finds that "on the basis of data involving 1,200 cases collected from four successive freshmen classes about equally divided between pledged and unpledged men, the pledged men contained fewer men of superior intelligence and more men of low intelligence." Other reports confirm this, but one report at least (Constance, 29) finds no difference between the two groups.

Time Spent in Study.—More intelligent students report spending less time in study than do less intelligent students. Spence (27-28) reports negative correlations of .29 and .46, and Crawford (29) a negative correlation of .20.

Sectioning.—Sectioning students on the basis of intelligence will not automatically lead to great gains. Burtt (23) found little difference in achievement in psychology in sections so formed. It is of no advantage if no change in instruction takes place. But if more work is expected of the brighter sections, the difference between sections immediately becomes marked.

Elimination.—Students scoring low in intelligence are eliminated from college earlier and in larger numbers than students scoring high. Rogers (26) found the average score in the Thorndike Test of those eliminated by the end of the second year to be 63.1 as contrasted with an average score of 67.2 for the entire group. Out of 292 freshmen Cuff (29) found that 22 per cent of those eliminated were below the 25th percentile, whereas only 17 per cent of those retained fell below this point. Ellis (26) finds the mean score of the eliminated to be 87.6 on the Miller test as contrasted with 94.1 for the retained, which latter score is + .28 S.D. higher. Jordan (25) gives the average Otis score for those eliminated as 143.6 as contrasted with an average of 153 for those remaining. All of these studies and many others show without a doubt that lack of intelligence is one of the basic causal factors for students leaving college whatever may be the reasons usually assigned.

Low and High Scores.—Many comparisons have been made between those scoring high and those scoring low on intelligence tests. Edgerton and Toops (29) find that the best ten per cent in intelligence earned 1.78 as many hours credit as did the poorest 10 per cent during the fifteen quarters of possible attendance. Similarly the best 10 per cent of students earned on the average 2.46 times as many points as did the poorest 10 per cent of students. They say, "A student who has a low intelligence test percentile, a poor high school record and entrance conditions has rather poor college prospects." And they find from their analysis that the university "is not securing the attainment from the superior student that is clearly possible." Erffmeyer (24) also compares the low and high intelligence scores and finds that 38 per cent of the low

and only 3 per cent of the high had conditions in their work; 64 per cent of the low and only 9 per cent of the high had failures; 74 per cent of the low and only 12 per cent of the high were reported as delinquent in their work. And Burwell and MacPhail (25) estimate that a freshman in the lowest psychological decile has only two chances out of five of remaining more than one year and only one chance out of five of graduating. Those who score below 100 on the Army Alpha are poor college risks, according to Jones (29). Among forty such students, he found that 82 per cent failed to maintain an average mark of C in their work.

Normal Schools and Teachers Colleges.—The intelligence of students in normal schools and teachers colleges is on the whole slightly below that of colleges in general. The college seems to attract and retain more students of high intelligence. There is great diversity in intelligence among normal schools themselves (Whitney, 23; Kirkpatrick, 22; Averill, 23; Madsen, 24; Sims, 29).

The correlations between academic work and intelligence are about the same as found in colleges, but the correlations between intelligence and estimates of teaching are very much lower (Pyle, 28; Thompson, 28; Frasier, 29; Morris, 29; Sorenson, 29; Broom, 29). These range from zero up to about .38. It is, of course, extremely difficult to estimate teaching success, and ratings of teaching ability in actual school work are generally extremely unreliable. Nevertheless, it is very obvious that the correlation between abstract intelligence and teaching ability is low. A given minimum of abstract intelligence is required of the teacher, but additional amounts above that do not seem to influence teaching ability. Other traits then become the determining factors. Intelligence tests in teacher training institutions function in the same way as they do in other colleges, but they do not aid specifically in predicting teaching success (Frasier and Heilman, 28).

Administrative Uses of Tests.—Intelligence tests are to-day being widely used in colleges and universities through-

out the country. Toops' (26) survey in 1923-24 showed that 60 per cent of 110 colleges were officially using tests and this percentage has probably increased greatly since that time. The uses of intelligence tests reported by various colleges may be listed as follows:

1. As a partial basis for admission.
2. In determining dismissal for low scholarship.
3. In determining probation for low scholarship.
4. In dealing with disciplinary cases.
5. In determining amount of school work to carry.
6. In determining amount of outside work for self-support.
7. To encourage bright students to undertake graduate work.
8. To motivate the bright student who is doing poor work.
9. To select assistants and student clerical help.
10. In making recommendations for scholarships and fellowships.
11. In determining membership in honorary scholastic societies.
12. In sectioning students in classes.
13. In notifying instructors where the scholarship is not up to expectation from the tests.
14. In admitting rehabilitation men who have not completed common school work.
15. In giving advice on study programs.
16. In the appraisal of transfer credits from little known institutions.
17. In the vocational guidance of students.
18. In experiments of all kinds.

No college uses intelligence tests as the sole basis of admission. Success in college and in after life depends upon much more than merely abstract intelligence as measured by the ordinary group test. What the educational guidance of college men involves is brought out admirably by Colvin (21, 22) in his description of the personnel work at Brown University. In addition to the intelligence tests, information is gathered as to a student's purpose and aim in life, his academic, vocational, non-academic and reading interests, his character qualities, his health, his family, his previous life, and so forth. Interviews

at the end of each college year are arranged, in addition to special consultations. The psychological test, therefore, is only one item, although an important item, in the total procedure. Colvin mentions the following ways in which intelligence tests have been found useful: (1) "They throw some light on the question as to whether a student is better fitted for a professional career or for business pursuits." Men with high intelligence ratings are urged to enter professional work, if their other qualifications are suitable. Society needs high grade men in the professions. (2) The tests indicate to some extent the type of mind that a student possesses. They seem somewhat unfair to men who are plodders, but careful and accurate thinkers. (3) They make possible a distinction between character qualities and mental alertness, e.g., cases of high intelligence ratings and low grades, and vice versa. (4) They throw light on the home environment and educational equipment of the student, especially in the case of those of foreign ancestry. (5) They show the presence or lack of scholarly ambitions and ideals. (6) They show whether it is desirable for a student to continue in college or withdraw. (7) They may serve as an incentive to work up to the level of one's abilities.

Summary.—College students are highly selected from the point of view of intelligence. Intelligence tests of all kinds have been tried out on them from the very beginning of the work in testing. The median scores on Army Alpha show a great range, indicating selective influences among the different colleges. All the medians are well above the median for unselected white men in the army. The correlations of intelligence with college marks vary greatly, the central tendency lying between .4 and .5. No particular course or subject seems to correlate uniformly higher with intelligence. Chronological age and time spent in study are both slightly negatively correlated with intelligence. In general the student who scores low earns fewer credits and much fewer academic honors than the student who scores high on the intelligence tests; he also is eliminated in much greater numbers and generally has difficulty

in graduating. In teacher training institutions the median intelligence seems to be slightly lower than in liberal arts colleges. Otherwise the relation of intelligence with other factors seems about the same as in colleges and universities. But there seems to be practically no relationship between teaching success and intelligence score. Seemingly a great number of colleges and universities now regard the intelligence test as an indispensable part of their machinery and the uses it is put to are manifold.

REFERENCES

Anderson, J. E. and Spencer, L. J. (26). "The Predictive Value of the Yale Classification Tests." *Sch. and Soc.*, 24, 305-312.

Averill, L. A. (23). "A Mental Survey of Fifty-nine Normal School Students: Some Correlations and Criticisms." *J. of Ed. Res.*, 7, 331-337.

Bell, J. C. (16). "Mental Tests and College Freshmen." *J. of Ed. Psych.*, 7, 381-399.

Bridges, J. W. (20). "The Correlation Between College Grades and the Alpha Intelligence Tests." *J. of Ed. Psych.*, 11, 361-367.

Bridges, J. W. (22). "The Value of Intelligence Tests in Universities." *Sch. and Soc.*, 15, 295-303.

Brigham, C. C. et al. (27). *Second Annual Report,* Commission on Scholastic Aptitude Tests. College Entrance Board, New York.

Brigham, C. C. (29). *Fourth Annual Report,* Commission on Scholastic Aptitude Tests. College Entrance Examination Board, New York.

Broom, M. E. (29). "The Predictive Value of Three Specified Factors for Success in Practice-Teaching." *Educ. Admin. and Super.*, 15, 25-29.

Burtt, H. E. et al. (22). "The University Intelligence Tests, 1919-1922." *Ohio State University Bulletin*, 27, No. 5.

Burtt, H. E. et al. (23). "Efficiency of Instruction in Unselected Sections in Elementary Psychology Compared with that in Sections Selected on the Basis of Intelligence Tests." *J. of Ed. Psych.*, 14, 154-161.

Burwell, W. R. and MacPhail, A. H. (25). "Some Practical Results of Psychological Testing at Brown University." *Sch. and Soc.*, 22, 48-56.

Caldwell, H. H. (19). "Adult Tests of the Stanford Revision Applied to College Students." *J. of Ed. Psych.*, 10, 477-488.

Calfee, M. (13). "College Freshmen and Four General Intelligence Tests." *J. of Ed. Psych.*, 4, 223-231.

Carothers, F. E. (21). "Psychological Examinations of College Students." *Arch. of Psych.*, No. 46.

Cattell, J. McK. (90). "Mental Tests and Measurements." *Mind.*, 15, 373-380.

Cattell, J. McK., and Farrand, L. (96). "Physical and Mental Measurements of Students of Columbia University." *Psych. Rev.*, 3, 618-647.

Cleeton, G. U. (27). "The Predictive Value of Certain Measures of Ability in College Freshmen." *J. of Ed. Res.*, 15, 357-370.

Colvin, S. S. (19). "Psychological Tests at Brown University." *Sch. and Soc.*, 10, 27-30.

Colvin, S. S. (21). "Educational Guidance and Tests in College." *J. of Appl. Psych.*, 5, 32-38.

Colvin, S. S. and MacPhail, A. H. (22). "The Value of Psychological Tests at Brown University." *Sch. and Soc.*, 16, 115-122.

Condit, P. M. (29). "The Prediction of Scholastic Success by Means of Classification Examinations." *J. of Ed. Res.*, 19, 331-335.

Constance, C. L. (29). "Greeks of the Campus." *Sch. and Soc.*, 30, 409-414.

Crane, E. (27). "The Results of Some Psychological Tests at Bryn Mawr College." *Sch. and Soc.*, 25, 640-644.

Crawford, A. B. (29). *Incentives to Study.* New Haven, Yale University Press.

Crawford, A. B. (30). "Forecasting Freshman Achievement." *Sch. and Soc.*, 31, 125-132.

Cuff, N. B. (29). "The Problem of Elimination from College." *Sch. and Soc.*, 30, 550-552.

Davidson, M. R. and MacPhail, A. H. (27-28). "Psychological Testing in a Women's College." *Person. J.*, 6, 266-275.

Davis, H. (21). "Army Alpha and Students' Grades; Illustrating the Value of the Regression Equation." *Sch. and Soc.*, 14, 223-227.

DeCamp, J. E. (21). "Studies in Mental Tests." *Sch. and Soc.*, 14, 254-258.

Dexter, E. S. (29). "Intelligence Test Score and Major Subject." *Sch. and Soc.*, 30, 779-780.

Downey, J. E. (17). "The Stanford Adult Intelligence Tests." *J. of Delinq.*, 2, 144-155.

Edgerton, H. A., and Toops, H. A. (29). *Academic Progress.* Ohio State University Studies. Contribs. to Administration, No. 1, O. S. U. Press.

Ellis, R. S. (26). "A Comparison of the Scores of College Freshmen and Seniors on Psychological Tests." *Sch. and Soc.*, 23, 310-312.

Erfmeyer, C. E. (24). "Intelligence Tests as an Aid in the Diagnosis of Academic Maladies." *Sch. and Soc.*, 20, 317-320.

Frasier, C. M. (29). "Intelligence as a Factor in Determining Student Teaching Success." *Educ. Admin. and Super.*, 15, 623-629.

Frasier, G. W. and Heilman, J. D. (28). "Experiments in Teachers College Administration: III. Intelligence Tests." *Educ. Admin. and Super.*, 14, 268-278.

Guiler, W. S. (27). "The Predictive Value of Group Intelligence Tests." *J. of Ed. Res.*, 16, 365-374.

Haggerty, M. E. (18). "Tests of Applicants for Admission to University of Minnesota Medical School." *J. of Ed. Psych.*, 9, 278-286.

Hammond, H. P. and Stoddard, G. D. (No date). *A Study of Placement Examinations.* University of Iowa, Studies in Educ., Vol. 4, No. 7.

Hollingworth, H. L. (16). *Vocational Psychology.* New York.

Johnson, J. B. (24). "Predicting Success or Failure in College at the Time of Entrance." *Sch. and Soc.,* 19, 772-776.

Jones, A. H. (29). "The Prognostic Value of the Low Range Army Alpha Scores." *J. of Ed. Psych.,* 20, 539-541.

Jordan, A. M. (25). "Student Mortality." *Sch. and Soc.,* 22, 821-824.

King, I. (17). "Relationship of Abilities in Certain Mental Tests to Ability as Estimated by Teachers." *Sch. and Soc.,* 5, 204-209.

King, I. and McCrory, J. (18). "Freshman Tests at the State University of Iowa." *J. of Ed. Psych.,* 9, 32-46.

Kirkpatrick, E. A. (22). "Intelligence Tests in Massachusetts Normal Schools." *Sch. and Soc.,* 15, 55-60.

Kitson, H. D. (17). "The Scientific Study of the College Student." *Psychol. Monographs,* Vol. XXIII, No. 1. Whole No. 89.

Kornhauser, A. W. (27). "Test and High School Records as Indicators of Success in an Undergraduate School of Business." *J. of Ed. Res.,* 16, 342-356.

MacPhail, A. H. (24). *The Intelligence of College Students.* Baltimore.

Madsen, I. N. (24). "The Intelligence of Normal School Students and its Relation to Learning and Teaching." *Ed. Admin. and Super.,* 10, 445-456.

Memoirs (21), The National Academy of Sciences, 15, Part III, Chap. 19.

Miller, W. S. (25). "An Objective Test in Educational Psychology." *J. of Ed. Psych.,* 16, 237-246.

Miner, J. B. (27-28). "Scholarship and Intelligence." *Person. J.,* 6, 113-118.

Morris, E. H. (29). *Personal Traits and Success in Teaching.* T. C. Contribs. to Educ., No. 342.

Nelson, M. J. and Denny, E. C. (27). "The Terman and Thurstone Group Tests as Criteria for Predicting College Success." *Sch. and Soc.,* 26, 501-502.

Noble, E. L. and Arps, G. F. (20). "University Students' Intelligence Ratings According to the Army Alpha Test." *Sch. and Soc.,* 11, 233-237.

Perrin, F. A. C. (24). "The Freshman Psychological Examination in the College of Arts and Sciences of the University of Texas: *University of Texas Bulletin,* No. 2445.

Pyle, W. H. (28). "The Relation between Intelligence and Teaching Success: A Supplementary Study." *Ed. Admin. and Super.,* 14, 257-267.

Rogers, A. L. (20). "Mental Tests as a Means of Selecting and Classifying College Students." *J. of Ed. Psych.,* 11, 181-192.

Rogers, A. L. (24). "What Women in Colleges of Liberal Arts Specialize In." *Sch. and Soc.,* 20, 700-702.

Rogers, A. L. (24-25). "Mental Tests for the Selection of University Students." *Brit. J. of Psych.,* 15, 405-415.

Rogers, A. L. (26). "The Causes of Elimination in Colleges of Liberal Arts for Women." *Ed. Admin. and Super.*, 12, 145-154.

Root, W. T. (23). "The Freshman Thorndike College Entrance Tests, First Semester Grades, Binet Tests." *J. of Appl. Psych.*, 7, 77-92.

Rowland, E. and Lowden, G. (16). "Report of Psychological Tests at Reed College." *J. of Exper. Psych.*, 1, 211-217.

Sims, V. M. (29). "A Comparison of Normal School Students with Certain Other Groups." *Sch. and Soc.*, 30, 68-70.

Sorenson, H. (29). "Why Teaching Success does not Correlate Highly with Measured Intelligence." *Ed. Admin. and Super.*, 15, 602-606.

Spence, R. B. (27-28). "Factors Related to College Achievement." *T. C. Record*, 29, 504-514.

Stone, C. L. (22). "The Significance of Alpha in Colleges." *J. of Ed. Psych.*, 13, 298-302.

Symonds, P. M. (24). "The Significance of Intelligence Tests in the University of Hawaii." *Sch. and Soc.*, 20, 601-606.

Terman, L. M. (Chairman) (23). *Report of Sub-Committee of Committee on Scholarship and Student Ability.* Stanford University.

Thompson, R. S. (28). "A Study of the Validity of the Downey Will-Temperament Test." *J. Ed. Psych.*, 19, 622-628.

Thorndike, E. L. (20). "Intelligence Examinations for College Entrance." *J. of Ed. Res.*, 1, 329-337.

Thorndike, E. L. (21). "On the New Plan of Admitting Students at Columbia University." *J. of Ed. Res.*, 4, 95-101.

Thurstone, L. L. (19). "Mental Tests for College Entrance." *J. of Ed. Psych.*, 10, 129-142.

Thurstone, L. L. (21). "A Cycle-Omnibus Intelligence Test for College Students." *J. of Ed. Res.*, 4, 265-278.

Toll, C. A. (28). "Scholastic Aptitude Tests in Amherst College." *Sch. and Soc.*, 28, 524-528.

Toops, H. A. (26). "The Status of University Intelligence Tests in 1923-24." *J. of Ed. Psych.*, 17, 23-36; 110-124.

Uhl, W. L. (19). "Mentality Tests for College Freshmen." *J. of Ed. Psych.*, 10, 13-28.

Van Wagenen, M. J. (20). "Some Results and Inferences Derived from the Use of the Army Tests at the University of Minnesota." *J. of Appl. Psych.*, 4, 59-72.

Viteles, M. S. (22). "A Comparison of Three Tests of 'General Intelligence.'" *J. of Appl. Psych.*, 6, 391-402.

Waugh, K. T. (16). "A New Mental Diagnosis of the College Student." *N. Y. Times Magazine*, Jan. 2.

Whinery, S. M. (26). "Psychological Test Ratings and College Entrance Age." *Sch. and Soc.*, 24, 370-372.

Whitney, F. L. (23). "Intelligence Levels Among State Normal School Graduates." *J. of Ed. Res.* 7, 229-235.

Wissler, C. (01). *The Correlation of Mental and Physical Traits.* Psych. Rev. Monograph, No. 16.

Wood, B. D. (23). *Measurement in Higher Education.* Yonkers.

CHAPTER XIII

THE SOLDIER

We have repeatedly referred to the intelligence testing conducted in the army during the years 1917 to 1919, and in this chapter we shall mention some of the important conclusions from that work. The detailed results of the whole work have been admirably presented in the "Memoirs" of the National Academy of Sciences, and we shall, therefore, make our statement very brief.

The title of this chapter, "The Soldier," must be understood to refer to the man in uniform during the period of the World War and not to the soldier in time of peace. The results are thus more characteristic of the general male population of the United States than they are of the professional soldier. As a matter of fact, the vast majority of the men tested were those inducted into the army by means of the Selective Service Act. We do not know whether the professional soldier of our ordinary peace-time army is the same as or different from the average drafted man. It is in the latter sense that we are using the term "soldier" in this chapter.

Because of the fact that the army in 1917-19 represented a fair sampling of the male population of the country, our data are of great importance. Up to the present time the army data represent the widest and most random sampling of adults that has ever been tested. To what extent the group of drafted men is selected it is impossible to say. We may note, however, several selective factors which were at work to cut off the population at the upper and lower ends of the supposedly normal curve of distribution of intelligence:

1. Superior men tended to become officers in the army.
2. Many superior men immediately volunteered their services and were not included in the draft.
3. Some superior men were undoubtedly excluded from the draft because of being employed in essential industries. Whether exemption for "essential industries" excluded more superior than inferior men it is difficult to tell, but probably the former is the case.
4. Many inferior men were excluded by the draft boards because of being mentally unfit for service.
5. Many inferior men were not included in the draft because of confinement in prisons, workhouses, reformatories, almshouses, and the like.
6. The most inferior men were in institutions for the feeble-minded.

The largest percentage, 51.1%, of exemptions or deferred classifications in the draft was given for "Dependency." It is doubtful whether this group can be considered superior to the drafted group. Lincoln (22) points out that in general the more intelligent man stays longer at his education and, therefore, is likely to marry later than the less intelligent. He is also generally economically more successful. Both of these factors would keep him from claiming exemption on account of dependency.

It may be, therefore, that these various selective forces, at work at both ends of the intelligence scale, roughly balanced each other and made the average drafted man a fair representative of the general population.

The Army Tests.—The work in the army extended from September, 1917, to January, 1919. Psychological testing was established in thirty-five camps and altogether 1,726,966 men were tested either by means of group or individual tests. This total includes 42,000 commissioned officers. Individual examinations to the number of 82,500 were given. The psychologists recommended 7,800 for discharge for mental defect, or 0.5 per

TABLE 13

EQUIVALENT SCORES OF TESTS USED IN THE ARMY

Alpha	*Beta*	*Point Scale*	*Complete Performance*	*Short Performance*	*Stanford-Binet*
...	...	...	...	...	4.0
...	...	...	3	2	4.5
...	...	...	4	5	5.0
...	...	...	6	8	5.5
...	...	...	9	12	6.0
...	...	31	17	17	6.5
...	...	36	30	24	7.0
...	2	42	41	33	7.5
...	6	46	52	42	8.0
2	11	51	62	53	8.5
4	17	55	72	67	9.0
7	24	60	91	86	9.5
11	30	64	114	108	10.0
16	37	68	135	127	10.5
21	42	71	153	144	11.0
27	47	74	166	158	11.5
33	53	77	175	169	12.0
40	58	79	183	179	12.5
47	63	81	189	188	13.0
56	67	83	195	197	13.5
63	71	85	201	205	14.0
71	75	87	208	214	14.5
78	78	88	216	223	15.0
85	81	90	223	232	15.5
93	84	92	230	241	16.0
102	88	95	237	250	16.5
114	91	98	244	259	17.0
125	95	100	251	267	17.5
137	99	...	258	275	18.0
147	104	...	268	283	18.5
161	108	...	290	291	19.0

cent of the total examined. They recommended 10,014 or 0.6 per cent for labor battalions because of low intelligence, and 9,487 or 0.6 per cent for assignment to development battalions for training and observation for possible use in the army.

TABLE 14

BASIS FOR THE ASSIGNMENT OF LETTER GRADES

	Alpha	*Beta*	*Point Scale*	*Complete Performance*	*Short Performance*	*Stanford-Binet*
A ...	135-212	100-118	Not given	260-311	275-308	18.0-19.5
B ...	105-134	90-99	95-100	240-259	250-274	16.5-17.9
C + .	75-104	80-89	90-94	215-239	220-249	15.0-16.4
C ...	45-75	65-79	80-89	190-214	190-219	13.0-14.9
C — .	25-44	45-64	70-79	150-189	145-189	11.0-12.9
D ...	15-24	20-44	60-69	90-149	85-144	9.5-10.9
D — .	0-14	0-19	0-59	0-89	0-84	0- 9.4

The tests used in the army were Alpha, a group test for literates; Beta, a group test for illiterates and foreigners; and individual tests such as the Stanford, Point Scale and Performance Scale.* The tests correlated well with various estimates of intelligence. "Alpha yields correlations with other measures of intelligence as follows: (1) with officers' ratings of their men .50 to .70; (2) with Stanford-Binet measurements, .80 to .90; (3) with Trabue B and C Completion tests combined, .72; (4) with examination Beta, .80; (5) with composite of Alpha, Beta and Stanford, .94; (6) in the case of school children Alpha correlates with teachers' ratings .67 to .82; school marks .50 to .60 . . . Beta correlates with Alpha .80; with Stanford .73; with composite of Alpha, Beta and Stanford .91" (Yoakum, 20).

The Average Intelligence.—If we ask what was the average intelligence of the drafted man, we may most readily

* For significance of letter ratings used on the group tests and for table of equivalent scores for the several tests used, see tables 13 and 14 on pages 318 and 319.

understand the result after the scores on the various tests have been converted into mental ages. These mental ages were equated with the Stanford Scale and they may, of course, be too low, if the standardization of the Stanford Scale was based upon children of better than average mentality. This has been done for a sampling of the different groups studied as follows:

	Mean M.A.	*No. of Cases*
White draft	13.1	93,965
Colored draft	10.4	18,892
White officers	17.3	15,544

The white draft is a fair sampling of the soldiers pro-rated by states as far as possible. The negro draft is also pro-rated by states with an additional group from the northern states.

We, thus, have the surprising and unexpected result of a mental age between thirteen and fourteen for the average white soldier. It is this result expressed in terms of mental age that has led to much popular discussion and misunderstanding. What it means is that the average drafted man does as well on the particular tests under discussion as the average thirteen- to fourteen-year-old child. This result was arrived at by equating the scores on Alpha with the mental ages obtained on the Stanford Scale by a group of some 653 soldiers, assuming that the standardization of the Stanford Scale for children is accurate and that thirteen-year-old ability on that scale represents what thirteen-year-olds in general can do. This is probably true, for we have in addition the work of Proctor (21) and Kohs working with children alone. On the basis of results from several thousand school children they make a score of 60 (the average for the soldiers) equal to a mental age of thirteen. Similarly Doll (19) tested 514 school children and found that a score of 60 on the Alpha was the average for children between ages 11 and 12. In rural schools Lufkin (21) found a score of 60 to lie between ages 12 and 13. In addition to this we have the evidence from examination *a* in the army, which was the group test given before Alpha was introduced. A comparison

of the median scores on this test for soldiers and children gives the following results:

	Median Score	*No. of Cases*
Soldiers	161	15,140
Children, C.A. 13	157	389
Children, C.A. 14	170	274
Children, M.A. 12	145	25
Children, M.A. 13	173	35

It would seem, therefore, that the kind of ability demanded by the army tests is such that it does not increase much, if at all, beyond the age of thirteen or fourteen in the average individual. This is only true in so far as the men tested in the army are taken to represent a general sampling of the population at large. The implications of this for intelligence testing have already been referred to in previous chapters.

Mental Age Distribution.—The army data also give a distribution of the white draft according to mental age as follows:

Mental Age	*Per Cent*
16 and over	14.0
15	9.3
14	13.6
13	15.8
12	17.0
11	12.7
10	7.6
9	4.7
8	3.15
7	1.75
6	0.2
5	0.1
4 and below	0.1

In so far as this represents the distribution of mental ages likely to be found in the country at large, it is important. The old definition of a moron, as having a mental age from twelve to nine, would result in some thirty to forty per cent of the

men being classed as morons or worse, and it is these results interpreted in this way that have led to very absurd statements in popular and semi-popular writings. As we have noted elsewhere, that conception of a moron was rapidly disappearing even before the army testing.

A study of the table shows us a much larger percentage of cases in the lower mental age groups than we might have suspected, and more evidence of this sort may lead us to still further revise our concept of feeblemindedness. On the suggested basis of an adult mental age of fourteen and an I.Q. of 70 or below signifying feeblemindedness, we should have all those testing below a mental age of 9 – 10 as probably feebleminded. This basis from the above table would give a percentage of almost ten as probably feebleminded, a percentage much higher than any estimate of feeblemindedness for the general population that has ever been suggested. If we take the more recent suggestions (see Chapter XIV) that only those with a mental age below 9 – 0 or below 8 – 6 be considered feebleminded, we obtain percentages of 5.3 or 3.7. In this connection it is interesting to notice the experience of the psychologists in the army camps and the working basis which they arrived at: "In general, subjects whose mental age is below eight should be seriously considered for discharge or development battalion. Those whose mental ages range from eight to ten should be considered for use in special service organizations or for assignment to development battalions." In 674 cases given individual examinations at a certain camp, practically all below a mental age of eight were recommended for discharge from the army. Only 12.4 per cent of the men given individual examinations who were recommended for regular service fell below a mental age of ten, and only 0.3 per cent below a mental age of eight. The suggestion here seems to be that a mentality of eight or less (comprising about 2.5 per cent) is likely to be of no use to the army. Those between eight and ten can probably be used, but may need more than average supervision. Perhaps this latter group would correspond to the moron in civil life,

as being a person who needs more or less help or supervision. In civil life he is more likely to fail and get into difficulty, because there is lacking the constant supervision given in the army. Indeed, it would seem that the army might very well make use of individuals of this level of mental ability for the many simple and routine tasks which have to be done, because the army provides just that degree of routine and supervision which is helpful and valuable for individuals of this level of mentality.

Again the evidence from a study of certain disciplinary cases in the army shows the inability of the man of lower mentality to make the proper adjustments in the long run. Of 479 cases studied, 66.4 per cent fell below a C rating (average rating) and only 10 per cent above it; 20.7 per cent of these cases were rated D — or E, as compared with 7.1 per cent for the white draft.

Age and Intelligence.—To the extent that the age of thirteen or fourteen represents the limit of growth for the abilities tested by the Alpha examination, we should not expect any increase in ability from one year to another after age thirteen or fourteen. An examination of the scores of men in the army from ages twenty-one to thirty shows no evidence of increase, but the data in reference to this question are very scanty.

The relation between age and intelligence among officers has been carefully studied. The median score seems to decrease steadily with age, as follows:

Age	*21-22*	*23-24*	*25-26*	*27-28*	*29-30*	*31-40*	*41-50*	*51-60*
Score	146	146	147	143	141	133	125	120
No. of Cases..	985	2,330	2,434	2,101	1,665	3,963	1,635	240

The correlations between age and intelligence range from — .01 to — .19, showing a slight tendency for the score to decrease with age. As to whether this is due to an actual loss in intelligence or a difference in the selection of the various age groups, it is impossible to say.

Rank and Intelligence.—"There is a high positive correlation between the rank and the intelligence of the enlisted men. The relationship among the officers of the various ranks is not so clear" ("Memoirs," 21, p. 855).

Type of Service and Intelligence.—As far as officers are concerned, there was a great difference in the intelligence of different branches of the service. Below are shown the percentages of officers rating A or B (the two highest ratings on the test) and also the median score on Alpha.

Branch	*Per Cent A and B*	*Median Score*
Engineers	97	162
Field Artillery	93	150
Sanitary Corps	90	151
Field Signal Battalion	88	149
Machine Gun Battalion	87	141
Infantry	85	140
Quartermaster	78	134
Medical	77	129
Dental	75	123
Veterinary	61	117

These differences are greater than any due to chance and suggest either an actual difference in the men attracted to various professions or the selective factors at work in various professions or selective factors at work in the choice of officers for the different branches of the service in the army. A more detailed study by Cobb and Yerkes (21) of the medical group has been made, because of its peculiarly low position in the group, which is contrary to the popular opinion as to the intelligence of men in that profession. The authors conclude that the method of selection of medical officers accounts for most of the differences as they were taken without training in officers' training schools, and thus, missed the selective factors operating there. "The Medical Corps obtained the services of the ablest as well as the weakest men of the profession." And

finally, "it is reasonably certain, then, that age and method of military selection are largely responsible for the relatively low intelligence of medical officers."

Education and Intelligence.—The amount of education an individual can acquire is determined among other things by his intelligence. From the important and lengthy analysis of the amount of education previously obtained by the soldiers in the army, we shall note only the increase in score with amount of schooling as shown by the white draft:

	Grades 0-4	*5-8*	*H. S.*	*College*	*Graduate*
Alpha	22	51	92	118	146
No. of Cases ..	4,253	33,424	10,715	3,131	97

Conclusion.—In this chapter we have noted merely a few of the more outstanding facts resulting from the intelligence testing in the army. This extensive and important piece of work was done under the pressure of war-time conditions. It was done for a very immediate practical purpose, and not for purely abstract scientific reasons. It demonstrated conclusively that intelligence tests could be of great value in the organization of men, and that they could be easily and economically applied. General intelligence, as tested by the army tests, is one of the factors contributing to success in the complex of military efficiency. Incidentally the results of the army testing have shed a great light upon the mental make-up of the population as a whole, for the soldier during the war was a fair representative of the citizen of this country.

For further details of the army testing, the reader must be referred to the original sources mentioned in the bibliography. We shall have occasion in several other chapters to make use of the data as they apply to other topics treated in this book.

REFERENCES

Cobb, M. V. and Yerkes, R. M. (21). "Intellectual and Educational Status of the Medical Profession as Represented in the U. S. Army." *Bulletin of Nat. Research Council,* Vol. I, part 8, No. 8.

Doll, E. A. (19). "The Average Mental Age of Adults." *J. of Appl. Psych.,* Vol. III, December, pp. 317-328.

Lincoln, E. A. (22). "The Mental Age of Adults." *J. of Ed. Research,* Vol. VI, No. 2, pp. 133-144.

Lufkin, H. M. (21). "Report of the Use of the Army Alpha Test in Rural Schools." *Sch. and Soc.,* Vol. XIII, Jan. 1, pp. 27-30.

Memoirs (21). National Academy of Sciences, Vol. XV.

Proctor, W. M. (21). "Psychological Tests and Guidance of High School Pupils. *J. of Ed. Research Monograph.* No. 1.

Yoakum, C. S., and Yerkes, R. M. (20). *Army Mental Tests.* Holt, N. Y.

CHAPTER XIV

THE FEEBLEMINDED

The study of the feebleminded as a special group dates back to long before the advent of the psychological test, and there are, of course, many phases of the problem of the feebleminded that are not psychological and have little or no interest to the psychologist as such. It is, nevertheless, true that during recent times psychology has made important contributions to our understanding of these children, and the application of intelligence tests has decidedly changed and modified our ideas with respect to them.

It used to be thought that the feebleminded were a very small and very distinct group of individuals, forming more or less a distinct species and differentiated from the normal by a very marked divergence in their intellectual and emotional life. They were considered as beings apart, radically different from the rest of humanity. By some authorities they were thought to be closely allied to the insane and, indeed, a clear distinction between the insane and the feebleminded is not yet universal in the medical profession. Although the cause of feeblemindedness was unknown, it was more or less treated as an incurable disease, and the feebleminded were considered the special care of the physician. At times there arose a belief in the possibility of cure, either through the agency of the physician or the educator. The importance of education was first emphasized by Seguin (06; 07) in France and later on in the United States. Seguin's belief in the possibilities of improvement by means of his physiological method was very great and amounted at times almost to a belief in the possibility of cure. His enthusiasm and aggressiveness resulted in untold good in the

care of these children, and his emphasis upon education modified for the better a great many of the ordinary educational methods then in vogue. Although Seguin's work was of decided importance for the care and education of the feebleminded, it did not contribute very much to our understanding of the condition, and it remained for Binet to give us a deeper insight into the psychology of these children and to show us clearly the relationship between the feebleminded and the normal mind.

Definitions of Feeblemindedness.—The earliest definitions of feeblemindedness are legal in nature and Blackstone's may stand as an example of these: "An idiot, or natural born fool, is one that hath no understanding from his nativity, and is, therefore, by law presumed never likely to attain any." Other definitions are distinctly medical in type, considering feeblemindedness as a sickness or as due to the incomplete development of the brain. Thus, Delboe says, "Even the condition of many stupid men is to be designated as sickness," and Willis considers the brain the seat of the sickness and specifies the corpus callosum and the cortex (Hollingworth, 20).

Sociological.—The best modern definition, not specifically influenced by the work in mental testing, is that formulated by the British Royal Commission on the Feebleminded in 1908: "A feebleminded person is one who is capable of earning a living under favorable circumstances, but is incapable, from mental defect existing from birth, or from an early age, (a) of competing on equal terms with his normal fellows: or (b) of managing himself and his affairs with ordinary prudence."

Following closely on this definition, Tredgold (22), one of the best medical authorities on feeblemindedness, says: "We may accordingly define amentia as a state of restricted potentiality for, or arrest of, cerebral development, in consequence of which the person affected is incapable at maturity of so adapting himself to his environment or to the requirements of the community as to maintain existence independently of supervision or external support."

The definition of the British Royal Commission is essentially sociological in character, because it makes social competency the criterion for judging the mental status. Tredgold practically adopts this sociological viewpoint, but adds to it a physiological explanation in the phrase "restricted potentiality for, or arrest of, cerebral development." This physiological defect is due to "an imperfect or arrested development of the cerebral neurones."

All these sociological definitions are merely practical guides as to who should be committed to special care or supervision. They all beg the real question, because they say that the inability is caused by mental defect. Thus, having an incompetent before us, we must decide whether his incompetency is due to mental defect or laziness or dementia or whatnot. Furthermore, competency in society depends upon the degree of complexity in any given social group. The reliance upon this criterion alone might mean that an individual might be called normal in a simple rural environment, but feebleminded if he moved to a complex urban environment. From the point of view of supervision and control, this is perfectly satisfactory, but it gives us no insight into what degree of mental defect constitutes feeblemindedness.

Psychological.—The definitions given above may stand as samples of the best type of modern definition that has not been influenced by the mental test. The advent of scales for measuring intelligence naturally gave rise to a type of definition that referred to achievement on the scale as a criterion of feeblemindedness, and this type of definition we may regard as psychological. Binet and Simon (16) were naturally the first to suggest that degrees of feeblemindedness could be best defined in terms of mental age, but they did not limit their definitions entirely to this. Thus they say: "An idiot is a person who is not able to communicate with his fellows by means of language. He does not talk at all and does not understand. He corresponds to the intelligence level of a normal child ranging from early infancy to two years of

age." The mental age limits for idiots are 0 to 2 years; for imbeciles 2 to 7 years; for morons, above 7 years.

This mental age type of definition is again reflected in the definitions adopted by the American Association for the Study of the Feebleminded in 1911:

1. The term feebleminded is used generically to include all degrees of mental defect due to arrested or imperfect mental development as the result of which the person so affected is incapable of competing on equal terms with his normal fellows or managing himself or his affairs with ordinary prudence.
2. Idiots are those so deeply defective that the mental development never exceeds that of a normal child of about two years.
3. Imbeciles are those whose development is higher than that of an idiot but does not exceed that of a normal child of about seven years.
4. Morons are those whose development is higher than that of an imbecile but does not exceed that of a normal child of about twelve years.

In a similar fashion we find Doll (17) defining feeblemindedness as "a condition of arrested development, specifically of the general intelligence, which limits the individual to mental capacity not exceeding that of twelve-year-old normal children."

These mental age definitions all refer to individuals who have reached mental maturity. For those still growing, the definition is usually in terms of I.Q. Hollingworth (20) says: "Idiots grade roughly from 0 to 20, imbeciles from 20 to 40, and morons from 40 to 70."

All these mental age and I.Q. definitions are predominantly psychological. They substitute "mental" development for the physiological concept of "cerebral" development. They lay stress upon general development as opposed to specific abilities. The sociological results of this arrested development are

secondary, the assumption being that an individual of such restricted mentality would fall below the required social standards.

Statistical.—A mathematical, statistical or percentage definition of feeblemindedness was suggested at the same time independently both by Miner (18), and by Pintner and Paterson (16). This type of definition is a direct outgrowth of the hypothesis of a normal distribution of intelligence in the population at large. If intelligence is distributed normally among the population at large and if feeblemindedness is to be regarded simply as the possession of a limited amount of intelligence, we might agree to regard as feebleminded a certain definite percentage at the lower end of our distribution curve. The percentage to be regarded as feebleminded will depend very largely upon whether we wish to limit the term to those conservatively termed feebleminded at the present time, or whether we wish to enlarge the concept and include within the term the highest grade cases about which there is now much difference of opinion.

Pintner and Paterson (16) suggested that the lowest three per cent and Miner (18) the lowest one-half per cent be considered feebleminded. Burt (21) also argues that "for immediate practical purposes the only satisfactory definition of mental deficiency is a percentage definition based on the amount of existing accommodation." This practical definition is admittedly based upon the ability of society to care for its intellectually weakest members. At present Burt suggests that in the county of London "the mentally defective child is to be defined as one who for intelligence ranks among the lowest 1½ per cent of the school population of the same age." This lowest 1½ per cent would include, according to Burt, all those falling below a deviation of minus 2 S.D., or an I.Q. of about 70.

These statistical definitions have not found favor among psychologists, physicians or social workers. Perhaps this is owing to the fact that we still think of feeblemindedness as

some definite entity, like insanity or measles or chickenpox, which an individual either has or has not. We do not yet think readily in terms of continuous variation. The adoption of such a statistical definition of the term "feeblemindedness" would immediately divest it of all its sociological implications. An individual would be considered feebleminded simply because he fell in the lowest x per cent as determined by intelligence tests. Because of this he would need a special type of education. His character, temperament and emotional stability would determine whether or not he might need to be segregated in some colony or institution or whether he might be allowed to remain in society either with or without supervision.

In other words, then, most, if not all, the cases at present grouped under the heading "feebleminded," would be divided into two groups: (1) feebleminded, i.e., those testing below a certain level on standard intelligence tests, such standard being determined by the lowest x per cent of the population at large; (2) psychopathic, i.e., those showing deviations in temperament or character such as is found among the lowest x per cent of the population measured by character and temperament scales, when such scales are available. At present the second group would be ill-defined, but the hope is that in the future psychological scales for the measurement of these traits will be available and that then individuals could be as accurately classified in this respect, as they are at present in respect to their intelligence.

This grouping of individuals would have no direct bearing upon social competency or incompetency, except in so far as social incompetency might result from either feeblemindedness or psychopathy. The socially incompetent form a much larger group than merely the feebleminded and the psychopathic, as we have defined these two groups. The socially incompetent group includes in addition the insane, many delinquents, and many others who are incompetent from purely environmental, accidental, or temporary reasons. And many who would be

diagnosed by this method as feebleminded or psychopathic, would be socially competent in a simple environment.

It has been objected that this statistical or percentage definition of feeblemindedness is a changing concept, in as much as it would differ from country to country and from century to century. This is true. There probably are differences in the degree of intelligence of the inhabitants of different countries, and in as much as this is true the standard for any one country would differ from that of another. This, however, would have no significance for the practical problem of feeblemindedness in any one country. Ultimately the upper limit of intelligence for the lowest x per cent of the whole population of the world might be determined, but the possibility of this being done is slight and its practical value negligible.

As to the second objection that it might change from century to century, it is obvious that this is based on an optimistic belief in the development of human intelligence. In so far as the evolution of the human race indicates such a development, there would then be a gradual increase in the absolute amount of intelligence of the lowest x per cent, but relatively they would always be the least intelligent of the group and the least able to compete with their fellows. Eugenic measures tending to restrict the increase of this lowest group might help to raise the average mentality of the whole group.

It is obvious from a discussion of all these definitions of feeblemindedness that the difference of opinion arises because of the fact that the standard definitions generally try to combine two different things which are highly but not perfectly correlated. These two variables are lack of intelligence (in the rather narrowly restricted psychological sense) and social incompetency. The first can be fairly well measured. The second is a vague concept. If these two variables could both be accurately measured and if they then correlated perfectly, everyone would be agreed as to what feeblemindedness is. This is true of the lowest grades of the feebleminded, namely

the idiots and the imbeciles. They are low in intelligence and also socially incompetent. But when we come to the larger group above these lower grades, we find those socially competent but fairly low in intelligence and also those socially incompetent but fairly well endowed intellectually. What we call these individuals will depend upon whether we stress "the social incompetency" part of our definition or "the degree of mental defect" part of the definition. The term "feeblemindedness" is thus ambiguous. It attempts to be both sociological and psychological at the same time. We need two separate terms. Term A would refer merely to mental defect and those so low in general intelligence would need a special type of education. Term B would refer to social competency and those falling below a certain level would need social assistance. Those below the line in both of these measures would constitute the inmates of our institutions for the feebleminded.

Diagnosis.—Much .of the discussion above leads us naturally to the question of diagnosis, for our concept or definition of feeblemindedness will determine very largely our diagnosis. And a diagnosis is usually made with an eye to treatment. Hence it will come about that certain examiners will give more weight to intellectual deficiency and others to social incompetency. But for almost all examiners at the present time the main instrument for diagnosis is usually some intelligence scale, most frequently the Binet-Simon Scale.

This scale was indeed the outcome of Binet's revolt against the subjective and casual methods employed in the diagnosis of mental deficiency in the past. Now in its revised and modified forms it represents our most useful diagnostic instrument. If a child's I.Q. falls below 70 or if an adult (i.e., anyone above age 14) fails to make a mental age of 10 or above, there is grave suspicion of mental defect. If the retardation is very great and if nothing of importance turns up in the case history, one thorough individual intelligence test may suffice. If, however, the retardation is near the upper limits mentioned, and

if there are any extenuating circumstances, then the examiner usually proceeds to further testing. He may give a different type of scale, such as the Pintner-Paterson Performance Scale or the Porteus Scale, or he may give other single tests (see Bronner, 27), or he may use a different revision of the Binet-Simon Scale a week or so later, in order to check up on his first examination. From the total result of this testing and from his case history, he will then come to some judgment.

In all cases it is well to have a thorough medical examination before a final diagnosis is made and certainly before recommendations for future treatment are made. Before individuals can be committed to institutions for the feebleminded a medical examination is usually compulsory. Some types of insanity, deafness or visual deficiency, extreme shyness, language disability may all cause an individual to do very badly on an intelligence examination, and this inability may not be primarily due to mental defect.

With children below the age of three or four the diagnosis of mental deficiency is more difficult, because our present intelligence scales are not so accurate for these low ages. The Kuhlmann Revision of the Binet and the Merrill-Palmer Scale are the best testing instruments here. For very young children, particularly during the first year of life, Gesell's Developmental Schedules are of major importance. Except in cases of extreme mental retardation the certainty of diagnosis in these early ages is not so great. The young child is developing so rapidly and has so much potentiality of development ahead of him, that most examiners are cautious and willing to give any case the benefit of the doubt. Nor in general is the need for diagnosis so pressing. Re-tests at six months' intervals will ordinarily show whether a child is developing normally or so slowly as to put him definitely in the feebleminded class.

In addition to the results of the intelligence examination, the examiner will gather further information about the case. A developmental history will throw light upon the general growth of the child. The ages at which the child began to walk and

to talk are important here. Severe diseases in childhood should be noted. A family history will seek to establish the mental status of related individuals. A great number of feebleminded relatives will be indicative of a generally poor intellectual stock. A school history, if the child is of school age, is important, because mental deficiency early shows itself by inability to make normal progress in school. Standard educational tests may be given to supplement the information received from the school. With older children or adults a vocational history may be important. Wherever there is a suspicion of complicating factors a sensory examination and a neurological examination may be necessary. The final diagnosis, the prognosis and the recommendations for treatment will be made in the light of all this information. A mere intelligence examination is not sufficient, although this may be the one most important item. The work should be done by an experienced clinical psychologist who is able to evaluate all of the many factors in each individual case.

The Prevalence of Feeblemindedness.—Before the advent of the intelligence test, estimates of the number of feebleminded in any country were very rare and for the most part based upon the number of cases confined in institutions for the feebleminded. Obviously these represented only a small proportion of the mentally feeble according to the newer view which arose from mental testing. Indeed, Kuhlmann (16-17) calculates that only about 4.5 per cent of the estimated number of feebleminded in the United States are taken care of in institutions for the feebleminded. A small percentage is to be found in penal and charitable institutions, a very large percentage is made up of children in school, but the largest percentage of all, namely, 51.6 per cent, is unaccounted for and represents the feebleminded at large among the adult population.

Table 15 gives several estimates of the percentage of feeblemindedness in various countries or sections of a country.

TABLE 15

Estimates of the Prevalence of Feeblemindedness

Authority	*Percentage*
British Royal Commission, 1908	0.40
Oregon Survey, Carlisle, 1921	0.50
United States, Bailey and Haber, 1920	0.65
Wisconsin Schools, Cary, 1916-18	0.70
British Mental Deficiency Committee, Report 1929	0.73
Oneida County, N. Y., Carlisle, 1918	0.73
Porter County, Ind., Clark, 1916	0.90
Rural County, Ohio, Sessions, 1918	1.80
Rural Survey, Del., Mullan, 1916	1.80
Toronto Schools, Smith, 1920	2.00
Goddard's estimate, 1914	2.00
Terman's estimate, 1916	2.00
Cleveland Survey, Mitchell, 1916	3.00
Popenoe's estimate for U. S. A., 1929	4.00
X County, Calif., Terman, 1918	4.24
Eight Minnesota towns, Kuhlmann, 1923	4.70
X County, Minnesota, Anderson, 1922	6.10

Some of these percentages are opinions, others are based upon the results of surveys of school children with intelligence tests and others upon rather detailed surveys of the total population of a given area. In the light of our discussion as to the definition and diagnosis of feeblemindedness, it can readily be seen that workers investigating the same area might easily differ as to the number of feebleminded. Much would depend upon where a particular worker placed the upper limit of feeblemindedness. Furthermore, a given section of the country might differ from another in the number of feebleminded, although hardly to the extent of having more than eight times the number of feebleminded as is the case with X County, Minn., when compared with Oneida County, N. Y. Popenoe's (29) recent estimate of 4 per cent for the United States is based upon the results of various surveys of

school systems, states and so forth. This would mean about five million feebleminded persons. This estimate is twice that of such authorities as Goddard and Terman. It is also very much higher than the estimate of Bailey and Haber (20), based upon the findings in the army during the war.

The two British estimates of 0.40 per cent in 1908 and 0.73 per cent in 1929 merit particular attention. The Royal Commission on the Care and Control of the Feebleminded was appointed in 1904, made a detailed survey of the number of persons suffering from all forms of mental disease as of January 1, 1906, and published its final report in 1908.* The method adopted was to select certain typical areas, urban, industrial, agricultural and mixed, and then to comb these areas very thoroughly for all feebleminded individuals. No intelligence tests were used. The medical men making the enquiry visited schools, penal institutions, asylums, hospitals, charitable institutions and gathered information from all medical practitioners, clergymen, police, charitable organizations and any other person or agency likely to know of mentally deficient individuals. From the number of feebleminded found in these selected areas, an estimate for the whole country was arrived at. The Mental Deficiency Committee of 1929 made a similar survey, with the notable addition of intelligence testing. Group tests were given to all retarded children and then individual Binet tests to all the suspected cases. Age fourteen was used in calculating I.Q's and the following ranges for the three degrees of amentia were adopted: idiots below I.Q. 20; imbeciles from 20 to 40; and feebleminded (or morons) from 40 to 60. The percentage of defective children in the six areas examined was found to be 0.85 per cent, the urban areas showing 0.67 and the rural 1.05 per cent. The estimated incidence for the total population of Great Britain was 0.73 per cent. The author of this report notes that this estimate is about twice that of the Royal Commission made

* This is the reason why it is referred to in the literature of our subject as the Royal Commission of 1904 or 1906 or 1908.

23 years previously. He believes that this increase in the estimate is due to three important factors: (1) a more thorough ascertainment of cases, due to the use of intelligence tests; (2) the increased longevity of the feebleminded, due to better health conditions; (3) a slight increase in the birth rate of feebleminded cases particularly in the rural districts. He calls attention to the fact that not only has the total estimate for the feebleminded increased, but that the percentage of idiots has also increased, and idiots would not have been overlooked in the previous survey.

These two British surveys represent the most reliable estimates of the conservative type that we possess. The earlier one undoubtedly underestimates the number of feebleminded; the recent one probably includes most individuals with an I.Q. below 60. For adults (i.e. individuals age 14 and above) the upper limit of feeblemindedness would be about M.A. 8 – 6. In all probability, therefore, this estimate of about one per cent of the population of countries like Great Britain or the United States as being feebleminded is to be understood in the sense that the degree of mental defect is so great as to make them practically useless in modern society. If we enlarge our concept of social competency, undoubtedly a larger number will fail. The percentage of feebleminded will vary from about one to three according to our concept as to what constitutes social competency and as to what degree of intelligence constitutes mental deficiency.

Classification.—Classification of the feebleminded is either according to degree of mental defect or according to clinical types. Classification according to degree is most useful educationally and is of chief interest to the psychologist. Classification according to clinical types is due largely to the medical workers in this field and is of chief interest to them.

As there are all degrees of lack of intelligence from the normal to the extreme idiot without a break, any number of divisions are theoretically possible. In actual practice three main divisions of the total feebleminded group are now gen-

erally recognised. In America these three groups are called idiots, imbeciles and morons. In England the last group is called feebleminded and the generic term for the total group is ament. The idiot is an individual of very low mental age (up to about M.A. 2) who cannot look after himself adequately and cannot protect himself from common physical dangers. The imbecile is low in mental age (from M.A. 3 to about 6 or 7), can rarely be taught to read and write, and can only do the simpler tasks of life. The moron is much higher in mental age (from M.A. 6 or 7 to M.A. 8 or 9), can be taught the rudiments of reading, writing and arithmetic, and can do many tasks in life under supervision, but fails in competition with normal individuals.

The mental age dividing lines between these three groups are mere arbitrary divisions, as they must be where we have a continuous distribution, and there is not yet any common consent as to just where they should be placed. If, for example, the most recent suggestion adopted in the British Mental Deficiency Report of 1929 be taken, then the M.A. of idiots would range from 0 to 3, the M.A. of imbeciles from 3 to 5 – 8 or 5 – 9 and the M.A. of morons from 5 – 10 to 8 – 6. This upper limit of 8½ is a long way from the upper limit of 12, proposed in the early days of mental testing. It is certainly much more in keeping with modern thought on the upper limits of feeblemindedness. Some would extend the M.A. to 9 – 11 or just below M.A. 10, this giving an I.Q. of 70, calculated on the basis of C.A. 14. But no authority would now-a-days contend that all those having mental ages of 10 and 11 be classified as feebleminded. If we apply these limits to the army results, we note that an upper limit of M.A. 8 – 6 would cut off about 3.7 per cent; an upper limit of M.A. 9 – 11 would include about 10 per cent. Similarly when applying these limits to a random sampling of 4,925 school children not including children in special classes, I find that 1.3 per cent fall below I.Q. 60 and 6.6 per cent below I.Q. 70. It is, therefore, probably wiser to consider the upper limit of

feeblemindedness as lying somewhere in the neighborhood of I.Q. 60 and M.A. 8 – 6.

Under clinical types the earlier writers described almost exclusively certain physical or clinical varieties of particular interest to the medical profession. It is natural that the bizarre and peculiar should attract attention, even if we disregard the morbid interest which such cases may excite. In addition some of these cases yield to specific medical treatment and are, therefore, of importance from a medical standpoint. With reference to other groups the close resemblance among the members of the group gives rise to the natural supposition of a specific causative factor, which, if discovered, may give a clue to the effective treatment of the case. These clinical varieties are important medically, but not psychologically. They make up a very small percentage of the cases in any institution, and a much smaller percentage of the total population of feebleminded. Lapage (11), after describing six clinical types, says that, "91 per cent of the feebleminded are not of any special type."

These physical types generally include the Cretin, the Mongol, the Microcephalic, the Hydrocephalic, and sometimes the Paralytic and the Epileptic. This classification is very illogical, because individuals belonging to several of the groups are not always feebleminded, as is notably the case with the paralytic and epileptic, and also to a lesser degree the hydrocephalic. Again the groups overlap, because we may have a hydrocephalic condition in a microcephalic idiot, and epilepsy may appear in many of the other types.

Psychologically, these clinical varieties are of no interest, because the various groups are not differentiated as to general intelligence or any other mental characteristic. Cretins may range all the way from dullness to idiocy; Mongolians from high grade imbecility or, perhaps, moronity to idiocy. Microcephalics are usually idiots, but this depends to some extent upon the size of the cranium we decide to call microcephalic. Hydrocephalics range from superior intelligence to

low grade idiocy, the amount of intelligence depending upon how severely affected the individual is. We find all grades of intelligence among the paralytic and the epileptic.

Idiots Savants.—Earlier writers on the feebleminded paid much attention to a description of idiots savants, that is, idiots or imbeciles who possessed an extraordinary talent in some special direction.

The descriptions of most of these cases are not very convincing to the modern psychologist, because they lack the quantitative measures of intelligence by means of which he could judge the performances described. To the old-fashioned observer, who thought of the idiot as a thing apart, belonging to a different species, it must have been a decided shock to find one who could recite pages of a school reading book from memory, or remember all the birthdays of the children in an institution or construct presentable objects out of wood, and so forth. To the psychologist, thinking in quantitative amounts of intelligence, it is no surprise to find individuals of mental ages from seven to ten doing these or similar things, when he considers what a mentality of seven to ten can accomplish plus the results to be obtained by much practice, and making allowance for the presence of special abilities in limited amount among the feebleminded just as we find special abilities among the normal. All of the so-called idiots savants that it has been the privilege of the writer to see personally could readily be explained in this fashion. Those that are described in the literature are harder to account for, because we lack definite quantitative statements as to their intelligence level. One suspects in many cases that they were not technically feebleminded at all, being rather psychopathic or mildly insane types such as are to be found in most institutions for the feebleminded. The one case that has been described by Tredgold (22) at length was certainly not feebleminded. This man was called "The Genius of Earlswood Asylum," and possessed remarkable skill in drawing and mechanical ability. He constructed elaborate models of ships

and carried them out in the greatest detail. We have no record of this individual's mental age. We know, however, that he was deaf and had a deaf brother, and that his parents were cousins. He never went to school and was, therefore, very uncouth and ignorant. In fact, he must have been like the average deaf child who has been allowed to run wild without training. He was sent to the institution at the age of fifteen and there he had the first opportunity to learn things. He was put to work in the carpenter's shop and as he made such excellent progress, he was given more and more opportunity to follow his interests. Tredgold himself says that he does not believe the man to have been "intrinsically defective." It would seem not. It would appear rather as if he were an untrained deaf child with more than average ability, all of which, owing to the circumstances and his own interest, was turned in one particular direction. It is somewhat misleading to group such cases under idiots savants. They are not idiots and they are far from being savants.

The writer has found only three cases described by psychologists where intelligence tests are reported. One of these with a phenomenal memory described by Jones (26) had an M.A. of 11 – 10 and an I.Q. of 85 (14 year basis), and therefore may be ruled out at once as certainly not feebleminded. Another described by Otis (25) also had a phenomenal memory for dates, names and places, but his M.A. on two Stanford-Binets is given as 12 – 11 and 13 – 0. This would mean an I.Q. of 92 or 93 (14 year basis). His M.A.'s on other tests were generally lower, from 5 – 6 on the Porteus to 11 – 8 on Language Completion tests. Again this would not be considered a feebleminded case. We are thus reduced to the last case described by Minogue (23), a 23-year-old boy with an M.A. ranging from 7 – 5 to 8 – 8 and an I.Q. from 53 – 62 (14 year basis). This case seems indubitably feebleminded. The boy is described as emotionally unstable. His special accomplishment is music, playing by sight and by ear. He plays very difficult pieces, and has a phenomenal memory.

His father's family is described as very musical. Here then we would seem to have a moron with a special talent for music, concentrating all the intelligence he possesses in this one direction and achieving thereby much above the ordinary. But on the whole we must conclude that authentic cases of idiots savants, that is feebleminded individuals with special talents, are very rare indeed.

Mental Development.—The mental development of the feebleminded is not only slower than that of normal individuals but it also ceases earlier and begins to decline earlier. Kuhlmann (21) studied these factors by repeated examinations of 639 cases over a period of ten years. He found that during a year 4.8 per cent gained more than 12 months' mental age; 68 per cent gained from 0 to 12 months' mental age; 11 per cent made no gain and 16 per cent actually lost in mental age. This tendency to deteriorate showed a marked relationship to degree of mental defect, the largest percentage of those deteriorating being idiots and the smallest morons. Similarly Bonnis (26) in France, reporting on repeated re-tests of about 300 feebleminded cases, also found a tendency for the I.Q. to decrease, and Moore (29-30) in re-tests of 51 idiots reported that 35 showed a decrease in I.Q., 9 an increase, and 7 remained the same. The median change in I.Q. was $-.063$ points per month.

As to qualitative differences in the mental make-up of feebleminded and normal children, many workers have suggested that such exist, but up to the present there seems to be no agreement on this point. It would seem rather that feebleminded children vary in their mental make-up qualitatively just as do normal children. The type of study made by Fox (27) may lead to the discovery of such differences if they exist. This worker compared normal and feebleminded children of the same mental age. By giving them four group tests he found the feebleminded to be weakest on sub-tests involving following directions, counting, relationships, com-

parisons and discrimination of size. They were best on substitution tests.

As to whether the feebleminded are more uneven in the several mental abilities than normal children has been a matter of dispute among several workers. Brown (26) summarizes these divergent opinions and his own work seems to show that the dull are no more uneven in mental ability than are the normal, just as De Voss (25) found that the superior are no more uneven than are normal children.

With reference to learning ability in general, the work of Woodrow (17) shows that feebleminded and normal children of the same mental age have similar learning curves, and the amount of transfer from the sorting of geometrical forms to sorting sticks and pegs and cancelling forms and letters is also the same for the two groups.

In summing up these studies of the mental characteristics of the feebleminded one gets the impression that taken as a group they are very much like normal children of the same mental age. The mental development of the feebleminded is slower, but it does not differ qualitatively from the mental development of normal children.

Physical Development.—This is not the place to discuss the general physical development of the feebleminded (see Hollingworth, 20; Pintner 31), except to note that the picture is one of retarded development, but not quite so retarded as the mental. Physical stigmata used to be much emphasized, but Hollingworth (20) says: *"A diagnosis of mental deficiency cannot be made on the basis of stigmata."* Binet (10) made an extensive study of the physical signs of intelligence. In addition to head measurements, he studied dental irregularities, strabismus, malformations of the ear and palate, facial asymmetry, speech defects. The hands and physiognomies of normal and abnormal children were compared. Describing this work Peterson (25) says, "No important relations were found, and Binet says in conclusion that nothing is as deceiving as the physical appearance of intelligence, and that it is

necessary to react consciously against our instinctive impressions in this regard."

Summary.—We see, therefore, that the quantitative conception of intelligence, which has arisen on the basis of intelligence tests, has very definitely influenced our ideas with reference to the feebleminded. The idea that feeblemindedness is simply a low degree of general intelligence has made our definition of feeblemindedness more accurate and concise. The introduction of intelligence tests has led to larger estimates of the number of feebleminded than were common previously, and a greater interest in the question as to the prevalence of mental deficiency. If these estimates differ, it is not to be wondered at, because there is necessarily a difference of opinion as to the dividing line between feeblemindedness and merely backwardness, and also because there are differences between the accuracy of the tests used and the accuracy of the examiners in using them. As to where the upper limit of feeblemindedness should be drawn, conservative opinion would put it at I.Q. 60 and M.A. 8 – 6 (using age 14 as a basis for all I.Q. calculations above that age). Many, however, would include all below an I.Q. of 70 and an M.A. of 9 – 11. There are few authorities now-a-days who would consider mentalities of 10 and 11 as definitely feebleminded. As to what percentage of the population is feebleminded, conservative opinion would say about one per cent for the whole country, with wide variations for particular regions. A more liberal opinion as to what constitutes feeblemindedness would easily raise this estimate to two or three per cent. Furthermore, mental tests have given us a classification of the feebleminded according to amounts of intelligence into the three groups of idiots, imbeciles and morons, a classification which is of much greater significance socially and educationally than one which emphasizes physical peculiarities. For the education and training of the feebleminded, the use of intelligence tests has been of special value. They have enabled us to group children of like mentality together. They

allow us to forecast the type of training that is likely to be of advantage to a child according to his intelligence level; and they prevent us from hoping for too much improvement in individuals of limited intelligence. It is no exaggeration to say that the introduction of intelligence tests has led to a greater advance in our understanding and training of the feebleminded than had been made since the early days of the Christian era when they were first brought together into asylums as objects of charitable care.

REFERENCES

Anderson, R. G. (22). "Methods and Results of Mental Surveys." *J. of Appl. Psych.*, 6, 1-28.

Bailey, P. and Haber, R. (20). "Mental Defect: Its Frequency and Causes in the United States." *Ment. Hygiene*, 4, 564-596.

Barr, M. W. (13). *Mental Defectives, Their History, Treatment and Training.* Philadelphia.

Binet, A. (10). "Les Signes Physiques de l'Intelligence Chez les Enfants." *Année Psychol.*, 16, 1-30.

Binet, A. and Simon, Th. (16). *The Development of Intelligence in Children.* Trans. by Kite, E. S., Vineland, N. J. Training School Publications.

Binet, A. and Simon, Th. (16). *The Intelligence of the Feebleminded.* Trans. by Kite, E. S., Vineland, N. J. Training School Publications.

Bonnis, L. (26). *Le Développement de l'Intelligence Chez les Arrièrés.* Vannes, Commelin.

British Royal Commission on the Care and Control of the Feebleminded. (08). *Report of the Commission.* Wyman and Sons, London, 8 vols.

Bronner, A. F. et al (27). *A Manual of Individual Mental Tests and Testing.* Boston.

Brousseau, K. and Brainerd, H. G. (28). *Mongolism.* Baltimore.

Brown, A. W. (26). *The Unevenness of the Abilities of Dull and of Bright Children.* T. C. Contribs. to Educ., No. 220.

Burt, C. (21). *Mental and Scholastic Tests.* King, London.

Carlisle, C. L. (18). "The Causes of Dependency." *Eugenics and Welfare Bulletin*, XV. State Board of Charities, N. Y.

Carlisle, C. L. (21). "Preliminary Statistical Report of the Oregon State Survey of Mental Defect, Delinquency, and Dependency." *Public Health Bulletin*, 112, Treasury Dept. U. S. Public Health Service.

Cary, C. P. (16-18). *The State and the Public Schools.* State Dept. of Public Instruction, Biennial Report, Madison, Wis.

Clark, T. et al. (16). "Rural School Sanitation." *U. S. Public Health Bulletin,* No. 77.

De Voss, J. C. (25). "Specialization of the Abilities of Gifted Children." *Genetic Studies of Genius,* 1, Chap. 12, Stanford University Press.

Doll, E. A. (17). *Clinical Studies in Feeblemindedness.* Boston.

Fox, E. J. (27). "The Diagnostic Value of Group Tests as Determined by the Qualitative Differences between Normal and Feebleminded Children." *J. of Appl. Psych.,* 11, 127-134.

Goddard, H. H. (14). *Feeblemindedness, its Causes and Consequences.* New York.

Goddard, H. H. (14). *School Training of Defective Children.* Yonkers.

Hollingworth, L. S. (20). *The Psychology of Subnormal Children.* New York.

Ireland, W. W. (00). *The Mental Affections of Children.* Second Edition. Churchill, London.

Jones, H. E. (26). "Phenomenal Memorizing as a 'Special Ability.'" *J. of Appl. Psych.,* 10, 367-377.

Kuhlmann, F. (16-17). "Distribution of the Feebleminded in Society." *J. of Crim. Law and Crim.,* 7, 205-218.

Kuhlmann, F. (21). "The Results of Repeated Mental Examinations of Six Hundred and Thirty-Nine Feebleminded over a Period of Ten Years." *J. of Appl. Psych.,* 5, 195-224.

Kuhlmann, F. (23). "A State Census of Mental Defectives." *J. of Delinq.,* 8, 247-262.

Lapage, C. P. (11). *Feeblemindedness in Children of School Age.* Manchester Univ. Press.

Miner, J. B. (18). *Deficiency and Delinquency.* Baltimore.

Minogue, B. M. (23). "A Case of Secondary Mental Deficiency with Musical Talent." *J. of Appl. Psych.,* 7, 349-357.

Mitchell, D. (16). *Schools and Classes for Exceptional Children.* Cleveland Education Survey.

Moore, L. (29-30). "Mental Growth of Low-Grade Feebleminded." *Training Sch. Bull.,* 26, 88-95.

Mullan, E. H. (16). *Mental Status of Rural School Children.* U. S. Public Health Service, Vol. XXXI, Part 2, 3174-3187.

Norsworthy, N. (06). "The Psychology of Mentally Deficient Children." *Arch. of Psych.,* No. 1.

Otis, F. E. (25). "Phenomenal Memory in its Bearing upon Various Mental Tests." *J. of Appl. Psych.,* 9, 311-318.

Peterson, J. (25). *Early Conceptions and Tests of Intelligence.* Yonkers.

Pintner, R. (31). Chapter on "Feeblemindedness." Handbook of Child Psychology. Clark Univ. Press.

Pintner, R. and Paterson, D. G. (16). "A Psychological Basis for the Diagnosis of Feeblemindedness." *J. of Crim. Law and Crim.,* 7, 32-55.

Popenoe, P. (29). "The Extent of Mental Disease and Defect in the American Population." *J. of Juv. Res.,* 13, 97-103.

Pressey, S. L. and Pressey, L. C. (26). *Mental Abnormality and Deficiency.* New York.

Report of the Mental Deficiency Committee. (29). London, H. M. Stationery Office.

Seguin, E. (06). *Traitement Moral, Hygiène et Education des Idiots.* Paris, Alcan. 1846. Reprinted 1906.

Seguin, E. (07). *Idiocy:* Its Diagnosis and Treatment by the Physiological Method. 1864. Revised 1866 under title, "Idiocy and Its Treatment by the Physiological Method." Reprinted by Teachers College, Columbia University.

Sessions, M. A. (18). *The Feebleminded in a Rural County of Ohio.* Publication No. 12, Ohio Board of Administration.

Smith, W. G. (20.) "Immigration and Defectives." *Canadian J. of Ment. Hyg.,* 2, 73.

Sollier, P. (01). *Psychologie de l'idiot et de l'imbecile.* Second Edition. Alcan, Paris.

Terman, L. M. (16). *The Measurement of Intelligence.* Boston.

Terman, L. M. et al. (18). *Surveys in Mental Deviation.* Calif. State Board of Charities and Corrections.

Tredgold, A. F. (22). *Mental Deficiency.* Fourth Edition. New York.

Woodrow, H. (17). "Practice and Transference in Normal and Feebleminded Children." *J. of Ed. Psych.,* 8, 85-96 and 151-165.

CHAPTER XV

THE SUPERIOR

The Discovery of Superior Intelligence.—The child of superior intelligence has been discovered by the intelligence test. Previous to this time we have had geniuses, peculiar freaks and extraordinary prodigies, and the connotation attaching to such words as "genius," "prodigy," "precocity," and the like, indicates that they were regarded as something apart, as something unhealthy and slightly abnormal. We had to wait for the intelligence test to give us a saner viewpoint, to give us a better definition of superior intelligence and to show us that superior intelligence is not nearly so uncommon as we seem to have imagined.

Interest in the child of superior intelligence has been much slower in developing than interest in the child of inferior intelligence. This is due to two reasons: (1) the limitations of the earlier intelligence scales, and (2) the practical problems arising in the school situation.

The original Binet Scale was not well adapted to the discovery of superior intelligence, except in younger children, and the uncertainty of the standardization of the early American adaptations made workers hesitate as to the reliability of their results. In addition to this the scales were frankly constructed by Binet for the diagnosis of inferior mentality. The object was not so much to find out what a child could do, but rather what he could not do. The great majority of children tested were cases suspected of mental deficiency, and, indeed, from this period there has come down to us the phrase "to submit a child to a mental test," carrying with it a suspicion of the integrity of his intelligence. To this same

attitude can be traced the name "nut test," and other similar slang phrases. If the early interest had centered on the superior child, if our first tests had been constructed especially to diagnose superiority of intelligence, then to be given a mental test would have been looked upon as an honor or privilege.

There is a very good reason, of course, why mental tests should have arisen first for the diagnosis of the feebleminded. This reason lies in the difficulties ever present in the schoolroom in handling the dull and feebleminded child. These children make their presence felt very decidedly because they simply cannot keep pace with the rest of the class, however much attention and instruction they may get. Something had to be done with them and so the school pressed for a solution, and the psychologist was stimulated in his attempts by the recognition of a very real and practical problem. On the other hand, the superior child has never been considered a problem in the schools, mainly because he has never really been recognized. He almost always can cover the required work, and, doing so, fulfills the main requirement of the school. If he is unruly or mischievous, the school can and does deal more or less adequately with this type of behavior, even although it does not recognize that it may sometimes be a symptom of superior intelligence. Again the school greatly resents the suggestion that it cannot recognize superior intelligence, that it is necessary to have psychological tests in order to discover it. Does it not daily and monthly and annually pile up a vast array of grades or marks, so that the sheep may be separated from the goats, so that it may reward its brightest scholars and admonish the laggards? In other words the school has in general tacitly assumed that amount of school work accomplished is a direct measure of general intelligence, and is only slowly beginning to realize the difference between educational attainment and general intelligence. Even to-day this distinction between knowledge and intelligence is not clearly in the consciousness of the

teacher. She is more apt to assume that, because a child has done good work in the class in which he happens to be, he is therefore of superior intelligence. And, conversely, if he does merely average or poor work, he is, therefore, of normal or subnormal intelligence.

We do not mean to leave the impression that before the general use of mental tests no attention had ever been paid to children of remarkable ability. We find many references in literature to especially bright children, and the biographies of many great men bear record of their superior performances in childhood. Nevertheless, such references are scattered and leave the impression of something peculiar and very uncommon. Superior intelligence has certainly not been recognized as a vital educational problem. It is becoming to be so regarded to-day, because of the scientific study of such children by means of intelligence tests.

Studies of Special Cases.—Before the use of mental tests the main interest in superior children centered very largely in the "prodigy," that is, the child of very exceptional attainments. In general the children of whom we have any record were made the object of educational experiments in order to see how much could be achieved by this or that particular method. In most cases the parents of the child participated actively in the work. The modern conception of inherited general ability is for the most part lacking, and the results obtained are attributed solely to the method of education followed. It will be of interest to mention a few of the outstanding cases.

In an old German book published in Goettingen in 1779, entitled "The Life, Doings, Travels, and Death of a very clever and very well-behaved four-year-old child, Christian Heinrich Heineken of Luebeck, described by his teacher, Christian von Schoeniech" (Berkhan, 10), we are given a description of one of the earliest prodigies on record. We are told that at ten months he learned the names of things on pictures. Before 12 months he had committed to memory the

best stories in the five books of Moses. At 14 months he knew the stories of the Old and New Testaments. At four years he could read, but not write. He could add, subtract, multiply and divide. He knew French and had learned 1,500 sayings in Latin. He learned Low German from his nurse and knew the most important facts about towns and places in geography. His fame spread over Europe and he was summoned before the King of Denmark. The little prodigy, however, was always sickly and he died at the age of four years and four months, "a wonder for all time," his biographer pathetically remarks.

In many respects our little Christian Heinrich corresponds very nicely to the popular notion of a prodigy, namely that he is sickly and perhaps nervous, that he has a prodigious memory and dies early.

Let us turn to another historical case which is more hopeful. At the beginning of the nineteenth century Karl Witte (14), a German pastor, undertook the education of his son, believing thoroughly in the importance of education in the first five or six years of a child's life. "Natural aptitudes" are not nearly so important as we imagine, according to Witte. His watchword seems to have been the statement of Helvetius, which he cites repeatedly, "Any man, normally well endowed, can become a great man, if he is properly educated."

Witte maintains stoutly that his boy was only of average intelligence. These are the accomplishments of the child. He learns to read before he is four years old and picks up writing shortly afterwards. At six he begins the study of French and then in succession Italian, Latin, English, and Greek. At seven years and ten months he demonstrates publicly in a school his ability to read Greek, Latin, Italian and French. At nine he knows as much as a youth of 18. (Would his I.Q. be 200?) He shows his fitness to matriculate at the University of Leipzig at nine years. He studies, when eleven, analytic geometry, and when twelve calculus at the Univer-

sity of Goettingen. Obtains the Ph.D. degree at fourteen and is made Doctor of Laws at sixteen.

Witte, the father, pursued no special method of education but spent a great deal of time with his boy. In spite of his statement that the child was only of mediocre ability, it would appear rather as if he had an intelligence quotient between 180 and 200. It is interesting to note that in a way Witte had our present-day conception of different degrees of ability distributed on a normal curve. He supposes that men's aptitudes are "capable of mensuration" and could be graded from one to 100 and continues, "the children who are born with aptitudes of 80, 85, 90, 95 and 100 are certainly as rare as those, thank Heaven, who are by nature step-motherly endowed with aptitudes of 25, 20, 15, 10, 5, and 1."

Karl, according to his father, remained a strong, healthy, playful child and never became vain and conceited. At 23 he was full professor of jurisprudence in the University of Breslau. Later he went to Halle where he remained teaching and writing for the rest of his life until death overtook him in full intellectual vigor at the age of 83. The dire prophecies of insanity and premature death that were launched at the father, because of "forcing" the young child's mind, were unfulfilled.

John Stuart Mill (Courtney, 89) is another of the infant prodigies of which we have a complete record. He was subjected to a strict and severe discipline by his father, who undertook his education from a very early age. He began Greek at three. "I faintly remember," he says, "going through Æsop's 'Fables,' the first Greek book which I read. The 'Anabasis,' which I remember better, was the second." Between the ages of three and eight he read many Greek authors, such as Herodotus, Lucian, Plato, and also in English such historians as Hume and Gibbon. It is comforting to learn that at this period the boy also read "Robinson Crusoe," "The Arabian Nights" and other lighter literature. Between eight and twelve he studied Latin, Greek and Mathematics. The list of

authors becomes formidable, including most of the well-known classical authors and shows a much wider range of reading than the average college undergraduate, who is studying the classics, is called upon to read. In Mathematics he studied geometry, algebra and differential calculus. He started to write a history of Roman government. He was interested in science and says, "I never remember being so wrapt up in any book as I was in Joyce's 'Scientific Dialogues.'" From twelve to thirteen he studied Logic and Political Economy. His father was harsh and irritable. Fear predominated in the education. He read many abstract things that he did not understand. "I never was a boy," he said, "never played at cricket. It is better to let Nature have her own way." It is evident that Mill was not happy, because of the temperament and attitude of his father. The system of education was not good, and none but a child of a very high intelligence quotient would have responded in the way Mill did. There are indications of "forcing," but the boy did not break down under it. On the contrary he lived to become one of the foremost of English logicians and political economists.

We have similar records of the early education of Macaulay, the two Thomson brothers and others. James Thomson at twelve and his brother William at ten became regular students at the University of Glasgow, and both took many prizes there. James became a great engineer and died at the age of 70. William Thomson, Lord Kelvin, became famous in physics and died at the age of 83.

Rather different from the foregoing is the case of Francis Galton. He is not usually considered as one of the young prodigies, but rather as an example of slowly maturing genius, inasmuch as his chief work was done in later life after the age of fifty. Terman (17) has given us a summary of his childhood achievements as described in Pearson's "Life, Letters and Labors of Galton." The little Galton was taught mainly by his sister, and was not subjected to any special

educational procedure or method as was the case with the other children previously described. His sister taught the little Francis his letters in play, so that he could point to them before he could speak. By the time he was five he could recite much of Scott's *Marmion*. When two and a half years old, he could read a simple book, and he could sign his name before three years. During his fourth year he wrote a short letter to his uncle, which has been preserved. The day before his fifth birthday, he writes to his sister: "My dear Adele, I am four years old and can read any English book. I can say all the Latin substantives and adjectives and active verbs besides 52 lines of Latin poetry. I can cast up any sum in addition and can multiply by 2, 3, 4, 5, 6, 7, 8, 10. I can also say the pence table. I can read French a little and I know the clock. Francis Galton. Febuary [sic] 15, 1827."

Here we have a summary of Galton's achievements at five in his own words and, as Terman points out, we have reason to believe that all of it is true. There is much more evidence of his rapid intellectual development at succeeding ages, as shown by his letters and the books he was reading. At fifteen he was admitted to the Birmingham hospital as a medical student. After studying this evidence of intellectual superiority, Terman concludes that Galton's intelligence quotient must have been nearly 200, and his contributions to science in later life would certainly support this conclusion.

The best source for the boyhood records of eminent men is the work of Cox (26). Here we have gathered together all the available material of the early years of three hundred eminent men, who were born between 1450 and 1850. From a study of these boyhood records, estimates of the probable I.Q.s of these men in childhood have been made. These estimates range from 100 to 200 with a mean I.Q. of 135. Men like Mill, Grotius, Goethe and Leibnitz are assigned the highest I.Q.s. It is of course obvious that much error may creep into an experiment of this sort, and the I.Q. assigned to any one individual is merely a rough estimate, depending

to some extent upon how much information about his boyhood years has come down to us. Nevertheless, the data as a whole show conclusively that these individuals were achieving in their childhood what is only achieved now by children of high I.Q.s. We can say with a great degree of certainty that eminent men in the past would have had high I.Q.s, if they had been tested in their childhood.

In addition to these cases of individuals now dead and of whom we have not intelligence ratings, there remain a few cases of very superior children who have been tested by means of intelligence tests and described in recent psychological literature. Garrison (17), Burke and Hollingworth report the case of E. who was 8 years and 4 months when tested, and obtained a mental age of 15 – 7, with an I.Q. of 187. In addition to this, E. did well on all other tests, except those involving manual dexterity. At the time of the test the boy was in the sixth grade. He was strong and healthy, but showed little inclination to indulge in games and sports. Both the parents showed high intellectual attainments.

In a report by the same writers (Hollingworth, 22) five years later we learn that E., when 12 years old, took the Thorndike college tests for entrance to Columbia College, and that he ranked second among 483 competitors, whose median age was about 18. On the Army Alpha he scored 194 and 201 on two different forms of the test. The boy entered Columbia College at the age of twelve, and, at the time the report was written, had attained a sophomore status. During his freshman year all his academic ratings were B or better, with the exception of physical education in which he received a C. He seemed to get along quite well with the other students and is described as a "good sport." In the five years during which E. has been studied, no tendency to become mediocre has appeared. During this time he has passed from 6th grade to 2nd year in college. "Average children, the country over, born when this child was born, and measuring

100 I.Q. when he measured 187 I.Q. are now in the 7th grade of the elementary schools" (Hollingworth, 22).

In a still later report of this same case by Hollingworth (27), we have an account of him at age 18. His score on the CAVD test is 441, which is about 4 P.E. above the average college graduate, or about in the top one quarter of one per cent of college graduates. He gained his B.A. degree and Phi Beta Kappa at age 15, his M.A. degree at age 16 and at the time the report was written had nearly finished his work for the Ph.D. degree. Up to this time he has certainly fulfilled all the expectations aroused by his I.Q. of 187 at the age of 8.

Coy (18) reports the case of a girl 9 years 10 months testing mentally 16 years 5 months with an I.Q. of 167 on the Stanford Revision. Other tests confirmed this mental age of 16. Her lack of schooling made her educational tests comparatively poor. In January 1917 she was in the fifth grade, but in June she was promoted to the seventh. She was thoroughly normal in all her social reactions, and indeed is described as being quite a "tom-boy." Both of the parents are characterized by high intellectual attainments.

Terman and Fenton (21) give a preliminary report of Betty Ford, chronological age 7 – 10, mental age 14 – 10, I.Q. 188 on the Stanford Revision. She ranks high in all other intelligence and educational tests. She shows a great interest and gift for the composition of poetry and prose, and much of her writing is considered to have intrinsic merit. The child has no known physical defects and always has been very healthy. Physically she is a year or so accelerated. Both parents are decidedly above average in intelligence.

Geniuses Who Were Dull in Childhood.—We cannot leave this topic of youthful precocity without calling attention to the popular conception that many dull children later become brilliant. This popular idea is supported by many stories of the boyhood of great men which tell of their poor standing in school. The list of such men is a long one (Swift, 08). Linnæus was advised to become a cobbler. Darwin was

considered below the common standard in intellect. Napoleon was mediocre in military school. Newton was the last boy in his class. Samuel Johnson was indolent. Swift was only allowed to take his degree in college by "special favor." Wordsworth was shiftless. It was doubted whether von Humboldt had the ordinary powers of intelligence. Heine, of course, was a failure, because he revolted at the formalism of school and college. And so it goes with Lowell, Goldsmith, Richard Wagner, Pasteur and Gladstone. And, finally, we have the remark attributed to Hume's mother, namely, "Our Davie's a fine good-natured cratur, but uncommon wake-minded."

It is a pity we do not have intelligence ratings of all these people in their childhood days. All the evidence of modern psychology is against the opinion that a child of inferior or even average intelligence can later become endowed with superior intelligence. In so far as the men cited above were men of superior intelligence, we must conclude that they were misunderstood as children, or they were not given the right kind of education, or because of environmental causes or character qualities they did not utilize all their intelligence in their school work. As we shall see, this latter factor is common among bright children in our schools to-day. Many a bright child learns little in school or, worse still, learns bad habits of study or learns to hate mental work, and the names cited above are probably historical examples of bright children who became valuable individuals in after life in spite of the handicaps of their education. History, of course, does not record the many bright children, misunderstood and wrongly educated, who never became famous in later life.

The Prevalence of Superior Intelligence.—If we raise the question as to the percentage of people of superior intelligence in the population at large, we find a different problem confronting us from the one as to the percentage of feebleminded. There has not previously existed any concept as to what is a person of superior intelligence, and, indeed, only

a vague conception of the more common word "genius." There has never been any attempt to survey the population or a part of the population in order to find out the number of gifted individuals as in the case of the feebleminded. Furthermore, the concept of superior intelligence, being so vague and ill-defined, has not had attached to it social significance. It is free and clear of sociological meanings, in contra-distinction to the concept of feeblemindedness, which the psychologist inherited burdened with social interpretations.

This being the case the psychologist has practically formulated his own definition of superior intelligence and he is doing this at the present time by considering the upper 20 or 25 per cent of individuals tested by means of reliable intelligence tests as superior. This upper group of 20 to 25 per cent is then divided into two or three groups, such as "superior," "very superior" or "genius," and the percentage of individuals in these groups becomes progressively smaller as we ascend the scale. We see here, therefore, an acceptance of the statistical type of definition, which has been rejected in the case of the feebleminded, because of the older meanings attaching to that word. Superior intelligence, therefore, implies the possession of a certain amount of intelligence, measurable by means of intelligence tests, and it does not imply the possession of any desirable (or undesirable) social traits, or of any specific ability to compete with decided success in the world at large. As defined, therefore, it is quantitative and psychological; it is not qualitative and sociological.

The point fixed by Terman on the Stanford-Binet Scale above which an individual may be considered of superior intelligence is an I.Q. of 110. The percentage of children having I.Q.'s greater than 110 in his investigation was twenty. Future investigations may show a larger or smaller percentage falling above this I.Q. point. At the present time an I.Q. of 110 is by common consent regarded as the dividing line between normality and superiority.

This division at 110 includes a great many cases, and many

studies have been made with more select groups, such as those above I.Q. 120 or I.Q. 130. About one per cent may lie above I.Q. 130 and these are sometimes called "the very superior" or "geniuses."

General Characteristics of Superior Children.—The popular conception of a youthful prodigy is that of a pale, sickly, nervous child, wearing glasses, avoiding the usual games of childhood, eventually becoming neurotic or insane, and, if not dying prematurely, at least developing into a useless member of society because of the over-cramming of his mind with useless knowledge. Although this conception is gradually weakening, it is nevertheless still common, and it evidently has been long prevalent. Over a century ago, Pastor Witte was warned of the dire consequences that would result, the insane asylum, and so forth, if he persisted in continuing his method of education with the youthful Karl. We find him, therefore, taking great care of the health of his son, and he presents evidence of his son's good health and interest in games and the out-of-doors.

The answer to such a picture of the gifted child is best found in the monumental work of Terman (25). About 1,000 children above I.Q. 130 were selected for study, and these are compared with children whose I.Q.s are normal. It was found that among the gifted the ratio of boys to girls was higher than that in the general population. In racial origin these California children were found to be mainly of Western European and Jewish stock. The Jewish stock contributed about twice that expected from the total Jewish population of the areas investigated. The average social status of the families was much higher than that of the average family. In general the family incomes are fair, and they live in superior neighborhoods, but there are isolated cases from very poor families living in inferior neighborhoods. These children come from families where there are distinguished relatives in much greater proportion than would be found in the average family. The vital statistics of the families show a healthier than average stock, with

few cases of insanity or feeblemindedness. The anthropometric measurements show the gifted group physically superior. The medical examinations show them also superior to average children. In school progress they are 14 per cent of their age above the norm in grade location, and 48 per cent of their age above the norm in intelligence, so that they are under-promoted to the extent of 34 per cent. Their school marks are better than those of ordinary children. On standard educational tests the E.Q.s of the gifted are high, but not as high as their I.Q.s. The gifted are no more uneven in their school abilities than ordinary children. Their occupational ambitions are higher than those of the control group. In general they have the same type of interests as ordinary children. They make more collections, particularly of a scientific nature. Their play interests are in general like those of the control group, with a somewhat greater interest in plays that require thinking. They are more mature in their play interests, showing a greater liking for quieter and less sociable games. These gifted children read a great deal more than does the average child. The average gifted child of 7 reads more books in two months than the average control child up to age 15, and the range of reading is much wider. In character and personality tests they are very superior, about 85 per cent of the gifted being above the median of the control group.

This work of Terman, briefly summarized above, is important because it gives us a rough picture of the gifted child and clears away the old ideas with reference to the sickly prodigy and the puny bookworm. What we note is the tendency for desirable traits to be positively correlated. Along with high intelligence goes general all-around superiority. Of course, there are individual exceptions. There are children of high I.Q. who are physically below normal, who do not like active games, who are nervous or vain and conceited just as there are normal or subnormal children with similar undesirable traits. But such undesirable traits do not tend to accompany superior intelligence. On the contrary they are less likely

to be found among children of high I.Q.s than among children of average or below-average I.Q.s.

Many other studies of gifted groups have been made, but no study gives us such a complete picture as that of Terman's described above. Jones (25-28) gave many tests to 120 children with I.Q.s above 130 and found them generally superior. They came from better-than-average homes and the group contained a large percentage of Jewish children. A comparison of two groups of high school seniors has been made by Yates (22). One group was composed of the 25 who scored highest on a combination of several intelligence tests, while the other group of 25 were those who scored around the median. The latter group are, therefore, not inferior in intelligence, but perhaps average high school seniors. If compared with the population at large, they would undoubtedly rank high. The differences between the two groups are interesting. The superior group get better academic marks; are rated higher in intelligence by their teachers; have more foreign parents; have fathers whose occupational status is higher; have the same type of homes in a material sense; have more "broken" homes (death, divorce, etc.); are not more likely to be "only" children; walk and talk earlier; are about equal in health; seem to mature earlier; have learned to read and write earlier; like school better; read much more widely. All of these conclusions are based on a very careful gathering together of the information required.

A study of 53 children with I.Qs. above 120 has been made by Root (21). The group surpasses normal children in their performance on about 19 psychological tests of all sorts of functions. In general the children come from good homes and the author lays much stress on the environmental influences.

Out of 776 high school students Almack and Almack (22) found 51 superior, i.e., having an I.Q. above 110. Thirty per cent of these were retarded in their school grade in comparison with their chronological age, and all, except three, were retarded in comparison with their mental age. Eight of them were rated "average" in intelligence by their teachers. The

median academic grade was "A." They exceeded the Smedley norms in height, weight and lung capacity.

A study of a somewhat different kind has been made by Duff (29). He attempted to follow the careers of all children with an I.Q. above 135 obtained during 1921-22 in a survey of about 15,000 children in Northumberland, England. There were 73 such children. He compares these with a control group of about 100 I.Q. of the same age and school. He finds a general superiority of the high I.Q.s to the controls. The intelligent group secure more education than do the controls. But seven of the controls were found in secondary schools and reports of these were all unsatisfactory. He compares these seven controls who are having the benefit of a secondary education with thirteen high I.Q.s who did not go to secondary school and he says, "we find that the Control are inferior to the Intelligent in spelling, and greatly inferior in extent and expression of their answers; their reading is less in quantity and inferior in quality; and their ambitions for their careers are lower. These are probably the points where their further education might be most expected to bring them above those whose education stopped sooner. It has not done so. The numbers concerned are tiny, but their indication is clear, that higher education cannot compare with innate intelligence as a differentiating force."

In Germany and England interest is also beginning to be aroused with reference to the child of superior intelligence. Psychological tests are being used to discover these children. Stern (20) reports briefly several investigations conducted in Germany for selecting children of superior intelligence. In Hamburg the problem was to choose 990 children for a special curriculum in the fifth year of the elementary school. In Berlin a special three-year high school course was organized for boys of superior intelligence. Several similar investigations are also described. Stern emphasizes the necessity for considering other factors besides the result of the intelligence examination. The decision should rest on (1) the teacher's judgment

of the pupil's ability; (2) the school record; (3) the teacher's psychogram of the pupil; (4) an educational examination; and (5) the intelligence test.

In England Ballard (22) reports the use of intelligence tests "to discover gifted children worthy of free secondary education among elementary schools of the county of Northumberland." Several educational committees in other parts of England seem to be employing tests for similar purposes.

Special Traits.—In addition to these studies attempting a general picture of the gifted child, we have several studies dealing with particular traits. Children with I.Q.s above 135 are not especially lacking in motor qualities according to Hollingworth and Monahan (26), for they found them to be superior to children of like C.A. in tapping. Again, the same authors (Monahan and Hollingworth, 27) compared such children with a control group of the same age and race, having a mean I.Q. of 100, on certain physical tests. They found no difference between the two groups on the standing broad jump. The average group did better on "chinning," but the superior did better on strength of grip.

Kiefer (29) compared children of high I.Q. with those of normal I.Q. on five motor tests and found no difference between the means for ages 9, 10 and 11. Some superior children are poor in motor development as Chassell (24) shows in her study of three such cases. But these are exceptional, and the evidence at present seems to point to a slight superiority, or at least an equality, of the child of high I.Q. on physical tests when compared with the child of average I.Q. and equal C.A.

With reference to the musical sensitivity of children of high I.Q., Hollingworth (26) finds them to be about up to the norm for children of like C.A. on the Seashore Music Tests.

Special Classes for Superior Children.—Some of the most important studies of superior children have been made for educational purposes and we shall describe briefly the results reported by some of the workers.

Whipple (19) during 1916-17 conducted an investigation as

to the mental tests which might prove most valuable for the selection of gifted children. For this purpose a class composed largely of superior children was formed and conducted as a special class for one year. Especially instructive in this report is the very conclusive evidence as to the inability of the teacher or principal to select the children of superior intelligence as defined by the psychologist. The thirty children for the Special Fifth and Special Sixth, as the Superior Class was called, were selected by the teachers. Considering an I.Q. of 110 and above as superior, there were 11 out of 30 below this line, the lowest of the I.Q.s being 99. And conversely, in the classes from which the superior children were selected, twenty children considered as average or below by their teachers were tested. Six of these proved to have I.Q.s above the median I.Q. of the Special Group. The child with the highest I.Q., namely 167, was not chosen by the teachers, but was discovered by means of the mental tests. The teacher, therefore, is likely to consider as superior, children who are merely normal, and to consider as merely average some children who are decidedly superior. Hence the imperative need of mental tests in selecting superior children. The average I.Q. of the Special Fifth was 119 with a range from 102 to 146; the average I.Q. of the Special Sixth was 116 with a range from 99 to 133. These two groups were, therefore, not entirely composed of the superior. Nevertheless, these two classes easily completed two years of the ordinary school curriculum in one year. Only nine failed to accomplish this double amount and of these nine cases, seven were not superior according to the criterion of mental tests. The superiority of this group, as a whole, is further demonstrated by Whipple by means of numerous mental and educational tests.

Coy (23) selected and supervised for two years a class for gifted children. All of the children composing this class were selected by intelligence tests. If reliance had been placed upon teachers' judgments solely, fifty per cent of the brightest children would have been ignored. The I.Qs. of the children ranged from 114 to 156 with an average of 129. The class covered

three years of the regular curriculum in two years and some of the children who started one half year behind covered three and a half years of the regular curriculum. In addition to the regular work, extra cultural studies, such as Greek History and French, were pursued, and the regular school studies were conducted on a broader basis. Standard educational tests were used to measure the progress at repeated intervals, so that the educational attainment was not based upon the subjective judgment of the teacher or investigator. At the end of this period they were promoted to an ordinary eighth grade class in which they more than held their own among children much older chronologically. The careers of these children have been studied for one year after they left the special class, that is, down to June, 1921. Their further progress, if followed, should be of great interest and value. Much of importance in this special class, which the writer himself watched, cannot be stated in quantitative terms. The spirit of the class was excellent and there was no problem of discipline. Children who had never known what hard mental work was, experienced it for the first time and learned their powers. Other children realized for the first time that they had serious competitors in the world, and this realization had a sobering and healthy effect upon their characters. Instead of fostering a spirit of "snobbishness" or "superiority," the special class brought a spirit of healthy, hard work, and eliminated any tendencies towards "snobbishness" which existed and which are more likely to flourish among superior children in the ordinary classroom where they run the risk, sometimes, of being held up as models.

The parents of the children were all very satisfied and most of them were loud in their praises as to what the special class had done for their children. To the writer this moral and character-building influence of the special class was particularly interesting, because it is from this angle that the policy of segregating superior children has been especially attacked. When we have adequate scales to measure moral and character

qualities, we shall be able to state such progress in quantitative terms.

In Detroit, special classes for bright children have been in operation since 1915 (Cleveland, 21). Since 1917, the children for such classes have been selected on the basis of mental tests. In 1921, there were three centers with over 200 pupils. The emphasis is placed upon a broader and richer curriculum, although the progress is also more rapid than in the ordinary class. The teachers of these special classes are enthusiastic about them, and we quote from one of them: "In my thirty-five-minute period with this eighth grade class, I accomplish as much as in a forty-five-minute period with ninth grade pupils. The comprehensive questions asked have in several cases surprised me—questions which would never occur to the average student. The problem of discipline does not exist. I have found an eagerness to learn, a deep sense of responsibility toward work assigned and a courtesy unparalleled in my teaching experience."

Children of superior intelligence are to be found in poor homes as well as in the homes of the well-to-do. Specht (19) describes a superior class formed in a public school on the lower East Side in New York City. The children were selected by means of intelligence tests and all of them had I.Q.s above 120. The median chronological age was 10.6, the median mental age 14.4, and the median I.Q. 137. During the first term they covered more than two grades of the regular school curriculum, and during the two subsequent terms they covered two grades each term. Many additional subjects were also studied. At the time the report was made, all of the children had entered high school and were doing well there.

Danielson (29) describes the classes for children of superior mental ability in Los Angeles. She points out that during 1928-29 fifteen special classes with 450 children were in existence during the first semester, and seventeen classes with 510 children during the second semester. But during this same period five times as many dull and feebleminded children were

provided with special instruction. The emphasis is still on the side of the dull child. In the special classes for superior children an enriched curriculum is provided. A distribution of the I.Q.s of 568 children in these special classes shows an average I.Q. of 134 with a range from 115 to 185. The distribution shows:

I.Q.	*n*	*Per Cent*
Above 140	132	23
120-140	412	72
110-120	24	4

The pupils in these special classes not only enjoy an enriched curriculum but they move more rapidly through the grades. Enrichment and acceleration inevitably go together. The degree of acceleration is shown by a comparison of the high school entrance ages of unselected and superior children:

Grade	*60,884 Unselected Pupils*	*85 Superior February, 1929*	*119 Superior June, 1929*
B 7	12-6	11-3	11-1
A 7	12-11	11-5	11-7
B 8	13-5	12-5	12-0
A 8	13-9	12-2	12-4
B 9	14.0	13-2	12-11

And in the high school Jones and McCall (26) note that children who have been accelerated in the grades earn 81.2 per cent promotions in subjects as contrasted with 77.6 per cent promotions for all pupils.

The problem of the young bright child and his first entrance into school has been studied by Lincoln (29). He studied 54 young bright children who had been admitted to Kindergarten or grade I before the regular age. When these children were in grades IV to VII they were more than holding their own, being on the average $+ .33Q$ above the median. Washburn and Raths (28) also followed for 6 years 36 young bright chil-

dren admitted early, and they found less retardation among them. They find this policy to be justified by its results.

Lamson (30) followed up thirty-six children with I.Q.s above 135 in the elementary school through their high school career and found that they still rated high in intelligence, namely, in the top centile of adults. She also found that their achievement was superior to high school children in general, although their C.A. was two years less. They took part in more extra-curricular activities, and appear not to have suffered in health by entering high school so young.

Summary.—The discovery of the bright or superior child as we now know him may be definitely attributed to the use of intelligence tests. He has emerged as a definite entity and he is defined quantitatively by means of intelligence tests. The characteristics of the superior child are now fairly well known, and we have discarded the old ideas of the weakling prodigy. It appears rather that high intelligence goes with other desirable traits. The child of high I.Q. is certainly not as high in these other traits as he is in intelligence, but the important point established is that he is not below the average child in any of them.

The discovery of the bright child has led to an immediate interest in his education and already in many progressive school systems special classes for bright children are taken as a matter of course. But there are many problems that still remain unsolved. At what stage in their school career is it most desirable to segregate such children? So far, very little has been done in the first grade or kindergarten. Might it not be most profitable to begin as early as possible, before the child has the opportunity of forming bad habits of learning? What is to be done in the small school where there are not enough bright children of similar mental age to form a class? In other words, how can we apply the knowledge that we now have about the education of superior children, so that their education may become a regular part of the elementary school system, whether in the large city or rural school? In the large

city system the possible advantages of a special building for superior children should seriously be considered, just as it has been found advantageous for feebleminded children. Such a building would allow a better classification into more homogeneous groups, and the curriculum for the whole building could be modified to meet any needs thought desirable.

Shall the emphasis be placed upon rapid promotion or upon a broader and richer curriculum? Do these two aims necessarily exclude each other? Shall we not find that both can readily be realized? What shall be done with the superior child who is ready for high school at an age much below the average? Is he able to take part in the social life of the high school, and if he is, is it injurious or advantageous to him?

Again these same problems will confront us with reference to college. Is the very young freshman missing anything vital in education, because of the discrepancy between his age and the age of the average freshman? If he is losing anything we must ask further whether this is more serious than being kept back in his mental development by the conventional rate of progress through school. Will these supposed drawbacks in high school and college be really serious if we have a fair number of younger bright children scattered through the high school or college? In the large city school system, we must also consider the advisability of setting aside one high school for children of superior intelligence. In such a school, the average chronological age will be much below the average of an ordinary high school. The social activities in such a school would conceivably be very different from those usually found in the present-day high school. Whether such a segregation would be desirable or not, only experiments can show.

At any rate, there is no doubt that our present system of dealing (or rather not dealing) with the superior child is decidedly wasteful. A great amount of good intelligence lacks opportunity for development, and the modern democratic state is greatly in need of all of it. It is axiomatic that such a democracy needs an intelligent citizenship, and hence the im-

portance of free universal education, but it is just as important for a democracy to develop the right kind of leaders. It must, therefore, help to develop to the utmost all the desirable capacities of all its citizens.

REFERENCES

Almack, J. C. and Almack, J. L. (22). "Administrative Problems Connected with Gifted Children." *Ed. Admin. and Super.*, 8, 129-136.

Ballard, P. B. (22). *Group Tests of Intelligence.* London.

Berkhan, O. (10). "Das Wunderkind Christian Heinrich Heineken." *Zeitsch. f. Kinderforschung,* 15, 225-229.

Chassell, L. M. (24). "Three Children of Superior Intelligence and Inferior Motor Achievement." *J. of Appl. Psych.,* 8, 128-139.

Cleveland, E. (21). "Some Further Studies of Gifted Children." *J. of Ed. Res.,* 4, 195-199.

Courtney, W. L. (89). *Life of John Stuart Mill.* London.

Cox, C. M. (26). "The Early Mental Traits of 300 Geniuses." *Genetic Studies of Genius,* Vol. II, Stanford Univ. Press.

Coy, G. L. (18). "The Mentality of a Gifted Child." *J. of Appl. Psych.,* 2, 299-307.

Coy, G. L. (23). *The Interests, Abilities and Achievements of a Special Class for Gifted Children.* T. C. Contribs. to Ed., No. 131.

Danielson, C. L. (29). "Special Classes for Children of Superior Mental Ability in the Elementary Schools." *Third Yearbook.* Psychol. and Educ. Division, Los Angeles City Schools, 23-43.

Duff, J. F. (29). "Children of High Intelligence, a Following-up Enquiry." *Brit. J. of Psych.,* 19, 413-438.

Garrison, C. G. et al. (17). "The Psychology of a Prodigious Child." *J. of Appl. Psych.,* 1, 101-110.

Goddard, H. H. (28). *School Training of Gifted Children.* Yonkers.

Hollingworth, L. S. et al. (22). "Subsequent History of E; Five Years After the Initial Report." *J. of Ed. Psych.,* 6, 205-210.

Hollingworth, L. S. (26). *Gifted Children: Their Nature and Nurture.* New York.

Hollingworth, L. S. (26). "Musical Sensitivity of Children Who Test above 135 I.Q. (Stanford-Binet)." *J. of Ed. Psych.,* 17, 95-109.

Hollingworth, L. S. (27). "Subsequent History of E; Ten Years After the Initial Report." *J. of Appl. Psych.,* 11, 385-390.

Hollingworth, L. S. and Monahan, J. E. (26). "Tapping Rate of Children Who Test Above 135 I.Q. (Stanford-Binet)." *J. of Ed. Psych.,* 17, 505-518.

Jones, A. M. (25-28). "An Analytical Study of One Hundred Twenty Superior Children." *Psych. Clin.,* 16, 19-76.

Jones, V. A. and McCall, W. A. (26). "Application of Two Techniques in Evaluating Some Policies of Dealing with Bright Children." *T. C. Record,* 27, 825-835.

Kiefer, F. A. (29). "Manual Motor Correlation in Superior Children." *J. Appl. Psych.,* 13, 357-371.

Lamson, E. E. (30). *A Study of Young Gifted Children in Senior High School.* T. C. Contribs. to Educ., No. 424.

Lincoln, E. A. (29). "The Later Performance of Underaged Children Admitted to School on the Basis of Mental Age." *J. of Ed. Res.,* 19, 22-30.

Monahan, J. E. and Hollingworth, L. S. (27). "Neuro-muscular Capacity of Children Who Test Above 135 I.Q. (Stanford-Binet)." *J. of Ed. Psych.,* 18, 88-96.

Peter, R., and Stern, W. (22). *Die Auslese Befähigter Volksschüler in Hamburg.* Barth, Leipzig.

Race, H. V. (18). "Study of a Class of Children of Superior Intelligence." *J. of Ed. Psych.,* 9, 91-98.

Root, W. T. (21). *A Socio-Psychological Study of Fifty-Three Super-Normal Children.* Psych. Rev. Monograph, No. 133.

Specht, L. F. (19). "A Terman Class in Public School 64, Manhattan." *Sch. and Soc.,* 9, 393-398.

Stedman, L. M. (24). *Education of Gifted Children.* Yonkers.

Stern, W. (20). *Die Intelligenz der Kinder und Jugendlichen.* Barth, Leipzig.

Swift, E. J. (08). *Mind in the Making.* New York.

Terman, L. M. (17). "The Intelligence Quotient of Francis Galton in Childhood." *Am. J. of Psych.,* 28, 209-215.

Terman, L. M. (25). "Mental and Physical Traits of a Thousand Gifted Children." *Genetic Studies of Genius,* Vol. I, Stanford Univ. Press.

Terman, L. M. and Fenton, J. C. (21). "Preliminary Report of a Gifted Juvenile Author." *J. of Appl. Psych.,* 5, 163-178.

Washburne, C. and Raths, L. (28). "Selection of Under-age Children for Entrance into School." *Ed. Admin. and Super.,* 14, 185-188.

Whipple, G. M. (19). *Classes for Gifted Children.* Public School Pub. Co., Bloomington, Illinois.

Witte, K. (14). *The Education of Karl Witte.* Trs. by L. Wiener. New York.

Yates, D. H. (22). "A Study of Some High School Seniors of Superior Intelligence." *J. of Ed. Res.,* Monog., No. 2.

CHAPTER XVI

THE DELINQUENT

Amount of Feeblemindedness.—The problem of the psychological make-up of the offender against society is one that has always attracted the psychologist. It has been felt that in general the criminal or delinquent must differ psychologically from the non-delinquent person, and it was hoped that such difference would help to explain to some extent the nature and causes of delinquency. About the time that intelligence tests were being introduced into this country, the belief in physical differences between the delinquent and non-delinquent, notably the anthropometric measurements of Lombroso and his followers, was on the wane. It was natural, therefore, that intelligence tests should be tried out on delinquent individuals. The first intelligence examinations were made on inmates of institutions for the feebleminded, but the workers in these institutions were naturally interested in the problems of delinquency, because such problems were continually being forced on their attention in connection with the past history of the feebleminded children.

Very early, therefore, in the history of intelligence testing, we find reports of such tests given to delinquents, and the study of the mentality of the delinquent has been extensively pursued ever since. These studies include delinquents of all types, those who are troublesome in school, those who are brought before a court, and finally the convicted delinquent detained in a reformatory, industrial school, jail, prison, workhouse or penitentiary.

The outstanding characteristic of the earlier reports was the very large percentage of feeblemindedness found among delin-

quents. Studies showing 60, 70 and even as high as 90 per cent of delinquents testing feebleminded were reported, and there was a tendency among many to accept these reports at their face value and assume that the all-important and, perhaps, only vital difference between the delinquent and non-delinquent was a difference in general intelligence. A simple solution of the vast mass of delinquency would thus be a more efficient and thorough-going segregation of the feebleminded in the general population. Delinquency, however, is not such a simple matter as to be explained by any one single cause, such as feeblemindedness, alcoholism, poverty, lack of education, or the like. This over-emphasis of the amount of feeblemindedness among delinquents was immediately questioned by many psychologists, and it was rightly pointed out that some of it was due to the faulty construction of the intelligence scales then available. This was notably the case with the Goddard Revision of the Binet Scale, which was much too hard at the upper end. It penalized the older children, those between ages twelve and eighteen, and it is just at these ages that we find most of the delinquents confined in boys' and girls' reformatories.

The amount of estimated feeblemindedness among delinquents can best be appreciated from a study of Tables 16 and 17. In these tables some of the investigations have been recorded. In most of the cases the writers used the Binet Scale in some form or other, although in some of the later reports group intelligence tests were used. The reports are listed in chronological order and glancing down the column it can be

TABLE 16

DELINQUENT CHILDREN

Author	*No. of Cases*	*Per Cent F.M.*
Hill & Goddard (11)	56	93
Gifford & Goddard (12)	100	66
Bridgman (13)	118	89
Otis (13)	172	75

Author	*No. of Cases*	*Per Cent F.M.*
Pintner (14)	100	46
Pyle (14)	240	66
Bronner (14)	505	9-11
Williams (15)	400	36
Hickman (15)	229	75
Haines (15)	200	62
Williams (15)	150	28
Haines (15) (16)	1,000	24-29
Kohs (15)	335	65
Crane (15)	809 (boys)	39
	386 (girls)	72
Healy & Bronner (16)	2,000	11
Ordahl (16)	341	40
Fernald (16)	124	20-24
Bowler (17)	75	45
Hall (17)	607	35
Kelley (17)	296	20
Ordahl (17)	33	45
Faber & Ritter (17)	110	14
Bridgman (18)	205	36
Ordahl & Ordahl (18)	432	23
Miner (18)	123	7
Williams (19)	470	30
Henry (21)	50	34
Richmond (21)	38	16
Anderson (21)	311	21
Mateer (21)	553	45
Healy (22)	1,212	7
Mathews (23)	341	29
Blanchard & Paynter (24)	500	28
Burt (25)	197	8
McCaulley (25)	100	16
Maris (26)	148	8
Bridges & Bridges (26)	140	15-20
Healy & Bronner (26)	4,000	13
Adler (28)	577	31
Oseretzky (28)	378	21
Adler (29)	369	35

TABLE 17

DELINQUENT ADULTS

Author	*No. of Cases*	*Per Cent F.M.*	*Type of Institution*
Rowland (13)	35	31	State Reformatory for Women
Mass. Report (14)	289	51	Prisons and Industrial Schools
Spaulding (15)	400	44	Prison
Rossy (15)	300	22	State Prison
McCord (15)	50	54	Prostitutes at large
Weidensall (16)	88	40	Reformatory
Fernald (17)	100	41	Reformatory
Haines (17)	100	20	Penitentiary
Gilliland (17)	100	33	Workhouse
Pintner and Toops (17)	132	29	Workhouse
Ordahl and Ordahl (17)	49	29	Penitentiary
Terman and Knollin (18)	155	17	Penitentiary
Doll (19)	50	16	State Prison

seen that, in general, there is a tendency to become slightly more conservative as to the percentage of feeblemindedness in the later reports. The median of the percentages of feeblemindedness for the children's reports is 30 and for the adults 31. By taking the median percentage of the earlier reports and the median of the later reports, we obtain some measure of the trend of opinion during the past two decades. For the children, the median of the first 21 percentages is 45, and of the last 21 percentages the median is about 21. For the adults, the median of the six earlier reports is 42, and for the seven later ones 29. The sweeping statements of the first mental examiners that 70 or 80 per cent of all delinquents are feebleminded is reduced to 20 or 30 in later reports, and the most conservative workers would put the percentage still lower. Miner's (18) careful analysis of the general results and his own personal studies lead him to this conclusion: "I see nothing in

the present evidence from mental tests to indicate that the frequency of mental deficients who might justly be sent to institutions from among the ordinary children who come before the juvenile courts of the country, would be over 10 per cent."

The variation of the percentages listed in Tables 16 and 17 is not by any means wholly due to the difference in examiners' opinions as to who should or should not be diagnosed as feebleminded. Much of it is due to the difference in the mental make-up of the groups examined. Included in the groups examined are delinquent and troublesome children in the public schools, children brought before the Juvenile Court, children convicted by the Court, children confined in various correctional institutions, individuals confined in prisons and penitentiaries, as well as one or two groups of recidivists. In general it has been found that the percentage of feeblemindedness increases as we go from the unconvicted to the convicted delinquents. The feebleminded individual is more likely to be caught and convicted than is the individual possessing a greater amount of intelligence. In the same way the individual with several convictions against him is on the average less intelligent than the individual with only one.

The Average I.Q. of Delinquents.—Although the estimates of the amount of feeblemindedness vary greatly, there is more general agreement on the low average I.Q. of delinquents. Here are average or median I.Q.s of different delinquent groups reported in several recent studies:

Author	*Median or Average I.Q.*	*n*	*Type*
Mathews (23)	81	341	Girls' Reformatory
Blanchard and Paynter (24)	80	500	Problem Children
Burt (25)	89	197	Boys and Girls
McCaulley (25)	85	100	Boys' Disciplinary School
Bridges and Bridges (26)	85	130	Boys' Reform School
Healy and Bronner (26)	90	4,000	Repeated Delinquents

Author	*Median or Average I.Q.*	*n*	*Type*
Merrill (26)	82	236	Juvenile Court
Root (no date)	76	1,916	Adult males
Bridges (27)	88	33	Girls
Sullivan (27)	90	353	Boys' Reformatory
Adler (28)	78	577	Boys' Reformatory

These averages range from 76 to 90. The average I.Q. of 76 is for adult penitentiary prisoners. All the rest are for juveniles. The highest averages of 90 are reported by Healy and Bronner for repeated delinquents brought before the juvenile court, and by Sullivan at the Whittier State School in California. At this institution the policy at that time was to reject the obviously feebleminded cases and send them to other institutions.

Burt's (25) percentage distribution of 197 cases according to I.Q. is as follows:

I.Q.	130-39	120-29	110-19	100-09	90-9	80-9	70-9	60-9	50-9
Per Cent	1.0	1.0	2.0	13.7	30.0	29.0	15.7	6.6	1.0

We may, therefore, sum up by saying that the distribution of delinquents is heavily weighted at the lower end. Although there may be great difference of opinion as to what percentage is feebleminded, there seems to be general agreement as to the fact that the average delinquent is mentally below the average non-delinquent on the usual abstract intelligence test.

Army Alpha Results.—Several surveys made by means of the Army Alpha Test have attempted a comparison between delinquents and the general population as represented in the army, and the results seem to show that the distribution of intelligence ratings of the delinquents is practically the same as that of the drafted men in the army.

Doll reports the results of the Army Alpha tests on 839 prisoners in the New Jersey State Prison. He compares their scores with the scores obtained from 6,541 white draft recruits

at Camp Dix, New Jersey. The average score for the prisoners is only slightly below the average of recruits at Camp Dix.

The reason for this discrepancy is the excessive number of negroes and foreigners in the prison group. He has eliminated the negroes in the army draft data. He, therefore, concludes—"In general, then, it may be said that when allowance is made for selective influences on the basis of nationality and color, the mental constitution of the prison, as a whole, corresponds very closely to the average intelligence of adult males of the state as a whole." Only a small percentage of prisoners (7 per cent) obtained scores equivalent to those of army officers as compared with the recruits (13 per cent). Superior intelligence would, then, seem to be less frequent among prisoners than in the general population.

Anderson's (21) data would seem to confirm Doll's findings, and they are interesting inasmuch as they are based upon younger delinquents between the ages of eight and nineteen. This fact is of importance because of the general belief in the greater incidence of feeblemindedness among juvenile as compared with adult delinquents. The distribution of the 197 girls as given by Anderson is as follows:

Rating	*Classification*	*Per Cent Delinquents*	*Per Cent 94,004 Drafted Men*
A	Very Superior	0.5	4.1
B	Superior	3.0	8.0
C +	High Average	14.7	15.2
C	Average	25.4	25.0
C —	Low Average	26.9	23.8
D	Inferior	17.3	17.0
D —	Very Inferior	12.2	7.1

The distribution for the delinquents is very much like the distribution for the drafted men, with the exception of a slightly larger percentage of very inferior, and a smaller percentage of superior. The percentage of delinquents rated A and B totals 3.5 as compared with 12.1 for the drafted men.

If we consider all testing D — as feebleminded, we have a percentage of 12.2 feebleminded. On the basis of the Yerkes-Bridges Scale, Anderson in the same article reports 20.9 per cent feebleminded. This large discrepancy suggests that much higher standards are being used for the individual than for the group tests.

Data gathered on delinquent soldiers ("Memoirs" 21) show on the one hand a distribution for Fort Leavenworth prisoners very similar to the general distribution of the white draft, while for prisoners in guard houses in various camps a large percentage of men of inferior intelligence. The percentage distribution of these two groups compared with the white draft is as follows:

	E, D —	*D*	*C —*	*C*	*C +*	*B*	*A*	*Total*
Leavenworth Prisoners	6.0	18.8	20.8	23.8	16.0	8.8	5.8	3,368
White Draft	7.1	17.0	23.8	25.0	15.2	8.0	4.1	94,004
Guard House Prisoners	20.6	25.5	21.6	18.9	8.3	3.4	2.1	1,004

The prisoners at Leavenworth were convicted on serious charges, while those in camp were convicted on minor charges. Low intelligence would seem to be a factor in less serious delinquencies. About 300 of the Leavenworth prisoners were conscientious objectors on religious or political grounds. The results show this group superior in intelligence to the white draft. Hence, to some extent the group at Leavenworth is not representative of prisoners in general.

The most extensive comparison by means of the Army Alpha has been made by Murchison (26). His percentage distribution of 3,942 native born white criminals as compared with 44,223 white drafted men is as follows:

Letter Rating	*Criminals*	*White Draft*
E	7.5	7.5
D	6.9	8.8
C —	17.8	21.4

Letter Rating	*Criminals*	*White Draft*
C	28.5	28.7
C +	22.8	18.8
B	11.4	9.7
A	5.3	5.1

Comparisons of criminals by states with state drafts give the same results. In general, therefore, the native white criminal is as intelligent as the drafted man. Murchison finds that the negro criminal tests somewhat above the negro army draft. The foreign-born criminal is below the foreign-born drafted man. Recidivists test higher than first offenders.

These studies of the intelligence of delinquents as compared with the intelligence of the drafted men in the army seem to indicate that we may have to revise still further our conception as to the amount of feeblemindedness among delinquents. Our estimate as to the general intelligence of the population at large has evidently been too high. As this estimate has decreased in amount, the difference between the intelligence of the delinquent and non-delinquent groups has diminished.

Superior Intelligence.—So far most studies have agreed in reporting a much smaller percentage of superior intelligence among delinquents than is supposed to occur in random non-delinquent groups. Even in the studies of Doll and Anderson by means of group tests where the difference between the delinquent and non-delinquent groups was small, what difference there was depended mainly on the discrepancy at the upper end of the distribution. Either the percentage of superior delinquents is very small, or else, and this is most often the case, the investigator has been mainly concerned with discovering the mentally deficient and has included the superior among those who are not deficient or among those who test normal or above. It is, therefore, impossible to say what percentage of delinquents have superior intelligence, and our best opinion would be that this percentage is very small as compared with the percentage of those having superior intelligence among the

non-delinquents, whatever may be our line of demarcation between average and superior intelligence.

Williams (16) has discussed this problem and finds that of 300 delinquent boys only 20, or 6½ per cent, had an I.Q. of 102 or above, as compared with Terman's estimate of 48 per cent for unselected school children. If, however, we define superiority as meaning an I.Q. above 110, we find only three of the 300 delinquent boys in the really superior group. This is one per cent. The group test results of Anderson and Doll previously mentioned report 3.5 and 7 per cent superior.

Some other results are as follows:

Author	*n*	*Per Cent Above Normal*
Mathews (23)	341	1.8
McCaulley (25)	100	2.0
Maris (26)	148	1.8
Adler (28)	577	4.6
Adler (29)	369	3.0

In general, therefore, it would not seem as if the percentage of above-normal intelligence among delinquents approaches anywhere near the percentage among non-delinquents. The popular opinion that a large number of criminals are individuals of great ability and intelligence, and that a great many crimes give evidence of having been conceived and executed by people of superior intelligence is certainly not confirmed by the results of intelligence testing.

Other Studies.—There is an elaborate early study by Goring (13) on the English convict that merits attention.

Goring's study is mainly anthropometrical and is directed against the opinions of the school of Lombroso and the belief in physical differences between the criminal and non-criminal groups. Exact measurements of a great many physical characteristics were made on a random sample of 3,000 English convicts. These results have no direct bearing upon our study of the intelligence of the delinquent, but Goring's main conclusions are worthy of note in passing. "We conclude that crimi-

nals are not physically differentiated because they are criminals, but because of differences in age, stature, intelligence, etc., etc., and of the different social classes from which they are drawn." After comparing criminal and non-criminal groups, he says, "our inevitable conclusion must be that there is no such thing as a physical criminal type." All the differences between criminals and non-criminals disappear when we equate for differences in stature and age. Criminals do differ from the general population in being somewhat shorter (average two inches) and slightly lighter. "These are the sole facts at the basis of criminal anthropology."

As to the mental differentiation of the criminal and non-criminal, Goring's study is of interest, although we feel that it suffers much from the fact that no objective tests were used. The men were classified into five groups: (1) intelligent; (2) fairly intelligent; (3) unintelligent; (4) weak-minded; (5) imbecile, upon the subjective opinion of the prison physician. Comparing the results of this classification with the British Royal Commission's estimate of 0.46 per cent feebleminded in the general population, Goring estimates the amount of feeblemindedness in the prison population as ranging from a minimum of 10 per cent to a maximum of 20 per cent. Taking the latter figure, he finds a coefficient of association between lack of intelligence and crime of .79, and concludes that "defective intelligence is one of the primal sources of crime in this country." He further remarks, "But probably the chief source of the high degree of relationship between weak-mindedness and crime resides in the fact that the criminal thing which we call criminality, and which leads to the perpetration of many, if not of most, anti-social offenses to-day, is not inherent wickedness, but natural stupidity." It is this native lack of intelligence, rather than environmental factors, that is a cause of crime, for "crime in this country is only to a trifling extent (if to any) the product of social inequality, of adverse environment, or of other manifestations of what may be comprehensively termed the force of circumstances." Finally, his general conclusion on

physical and mental differences is, "that the one significant physical association with criminality is a generally defective physique; and that the one vital mental constitutional factor in the etiology of crime is defective intelligence."

We have quoted at length from Goring because of the exactness and care of the author, the wide range of the study and the importance of his final conclusions. Equally exact and careful is the work of Fernald, Hayes and Dawley (20) on women delinquents in this country. It differs from Goring's study in that estimates of intelligence were based upon mental tests. The Binet, Yerkes, Stanford and Woolley Scales were used as well as a group of performance tests. An elaborate social study was made of each case. The measurement of the intelligence of the cases is unquestionably much more accurate than Goring's measurement, based on subjective opinion. No percentage of feeblemindedness is given with which to compare Goring's estimate of 10 to 20 per cent. The average mental age of 447 delinquents is 11.8 as compared with an average M.A. of 13.4 for an unselected group of 653 army adults. The overlapping of the delinquent and non-delinquent groups is very great, and the data indicate a slighter degree of difference than Goring's data indicate. Furthermore, the writers are inclined to stress the importance of "the force of circumstances" in their final conclusions in which they mention "two lines of influence which seem to have a bearing on the problem of delinquency among women, namely: (1) poor economic background with few advantages or opportunities, and (2) a somewhat inferior mentality." And again, "We disagree with Goring in the preëminence attached to such a constitutional factor as defective intelligence in contrast with economic factors."

These two detailed and careful studies both agree in finding the delinquent less intelligent than the non-delinquent, but differ as to the amount. It is interesting to note that the study based upon objective tests is the one that finds less feeblemindedness. On the other hand, Goring's subjective estimates undoubtedly include cases of mental derangement and pos-

sible character deficiencies which might not be included in those cases in the other study testing mentally deficient, and this may account for some of the difference. The radical difference between the two studies is the contrasting emphasis placed upon the influence of environmental factors. It is a difference that is frequently met with in the interpretation of psychological and social data, and in general, the tendency of the psychologist is to emphasize the importance of original nature as opposed to environmental factors.

Other Types of Intelligence.—So far almost all the studies quoted have depended chiefly upon the use of verbal intelligence tests. We may say, therefore, that the delinquent possesses less abstract intelligence than the non-delinquent. Does he also possess less concrete or mechanical intelligence? Non-language group tests have been used by Slawson (26) and Hamill (23). Slawson finds delinquent boys somewhat better on the Thorndike Non-Language Test than on the National Intelligence Test. Hamill makes no comparison between tests, but finds that 1,340 misdemeanants make a much lower median score than do 50 patrolmen on the Pintner Non-Language Test. And Aden (26) finds her delinquents below the norms on the Pintner-Paterson Performance Scale. So far as this type of non-verbal or concrete intelligence is concerned the delinquent seems again to fall below the non-delinquent, but whether as far below as on abstract verbal tests, it is difficult to tell.

Turning now to results with the Stenquist Mechanical and Assembly Tests, we have three reports which find the delinquents up to or above the norms for ordinary school children. Dougherty (26) finds his cases to score above the norms for New York City boys; Slawson (26) says they are up to the norms; and Asher (27) reports median T-scores of 59 and 55. Dougherty suggests that some of the social maladjustments of our delinquents may be caused by too much stress on the literary (verbal) type of ability in school.

Sex Differences.—Several writers have commented upon a probable sex difference among delinquents with regard to intel-

ligence. Comparing similar institutions for men and women, it is generally found that the average mentality of the females is less than that of the males. Taking also into consideration the well known fact that there is a much greater proportion of delinquent males than females, and that courts and juries are more reluctant to convict the female offender than the male, because of the greater opprobrium which thereby rests upon the girl or woman, it is not surprising that institutions for female offenders should show a larger percentage of feeblemindedness than similar institutions for male offenders. The female offender, more particularly the girl, comes to the institution with a longer record of delinquency behind her than the boy. She has been given more chances; she is put on probation more frequently by the court; so that among those who fail to make good, we are in the long run more likely to find a large proportion of feebleminded girls.

Type of Crime.—The relation between intelligence and the various types of delinquency has been studied by many investigators. Nothing very definite has so far resulted. Almost every type of crime or misdemeanor seems to be represented by all grades of intelligence. There is, however, a tendency for vagrancy, drunkenness, assault and battery, begging, and the like, to show a lower average mental age than forgery, embezzlement, and allied crimes. Burglary, larceny, thieving, and the like, are represented by all grades of intelligence, and this is also the case with homicide. Particularly brutal murders seem often to be committed by feebleminded, epileptic or insane individuals. Women and girls convicted of sexual immorality in general test lower than those convicted of other offenses. Most investigators seem to agree that truancy among school children is closely related to delinquency. This relationship between truancy and delinquency has been discussed by Abbott and Breckenridge, who report that 40 per cent of 456 truants were not normal mentally. Further, Doll (21) is of the opinion that a large proportion, probably about two-thirds of juvenile delinquency is traceable to truancy. The

juvenile delinquent very often starts his career by being a truant, although, of course, it does not follow that all truants later on become delinquents. Obviously truancy is a symptom of some sort of maladjustment between the child and the school. The inability of the child to adjust himself adequately to the school environment may later on show itself in inability to adjust himself to his out-of-school environment, and so lead directly to anti-social conduct.

Merrill (26) gives the following median I.Q.s for juvenile offenders ranked according to type of crime:

Offense	*Median I.Q.*	*n*
Truancy	90.3	25
Larceny	84.5	20
Burglary	84.5	20
Vagrancy	84.5	8
Incorrigibility	83.2	49
Stealing	82.8	99
Immorality	77.8	36
Drunkenness	77.0	4
Forgery	74.5	10

There are two similar rankings of crimes for adult prisoners. The one below by Root is based on a survey of 1916 penitentiary prisoners:

Crime	*Median I.Q.*
Embezzling	103.75
Robbery	84.3
Forgery	83.75
Burglary	81.75
Larceny	78.3
Pandering	75.0
Arson	75.0
Rape	72.8
Sodomy	72.1
Homicide	70.9
Felonious assault	68.3

Crime	*Median I.Q.*
All predatory cases	80.3
All crimes	76.2
All sex cases	72.8
All violence cases	70.2

Murchison (26) ranks the crimes according to the percentage of superior (having a letter rating above C) and the percentage of inferior (having a letter rating below C) intellects:

Crime	*Per Cent Inferior*	*Per Cent Superior*
Fraud	22	52.9
Force	30.6	40.5
Thievery	31.8	40.7
Statutory	31	34.7
Physical injury	36.9	35
Dereliction	43.1	35.3
Sex	47.6	26.3

There is a certain degree of agreement between the results for adults reported by Root and by Murchison. Predatory crimes tend to be high and sex crimes low.

Future Success of Delinquents.—If the modern justification for our correctional institutions is the reformation of the delinquent, then the after-success of the inmate will be a measure of the efficiency of our system. To what extent this is true is unknown, as there are few, if any, satisfactory studies of this large and important problem. Only one aspect of the problem, however, concerns us here, namely, the relationship between the after-success of the delinquent and his intelligence.

Healy and Bronner (26; 29) have made the most significant studies of the results of different methods of handling delinquent children. They point out the great difference between the work in Boston and in Chicago. They show that successful adjustment is to some extent dependent upon type of mentality. Thus they report in one study 51 per cent successes for normal cases; 33 per cent successes for feebleminded; and

33 per cent successes for psychopathic. Those having I.Q.s above 110 show 66 per cent successes. In another study, discussing the value of the foster home, they find that 85 per cent of delinquents of normal mentality and personality were successful in foster homes, whereas only 40 per cent of delinquents of abnormal mentality and personality were successful in foster homes. Pintner and Reamer (18) tried to estimate the after-success of 26 delinquent girls ranging in C.M.A. from 70 to 109. The correlation between intelligence and their success, as estimated by the combined judgment of three observers, was + .16. They conclude that the mental tests are not prognostic of after-success. Further, they say, that those of poor intelligence seem just as likely to make good as those of normal intelligence. Clark (20) in a more detailed report of 223 delinquent boys concludes that "there is a distinct tendency for the boys of higher intelligence to have a better record of success than those of lower mentality." He divides the cases into six industrial groups and finds radical differences in the correlations between intelligence and success in these various groups. The coefficients range from + .74 for the agriculture, forestry and animal husbandry group to — .51 for the transportation group. In general, he finds that "a positive general relationship between intelligence and success record for the whole group was indicated by a coefficient of correlation of + .19." Note that this is about the same as that reported by Pintner and Reamer.

Other Factors Besides Intelligence.—It is the main purpose of this chapter to show the results of the intelligence testing of delinquents. This we have attempted to do. We cannot, however, refrain from mentioning other mental factors that are of undoubted importance in the study of delinquency, so that we may not lose sight of the fact that the problem of adequately understanding the delinquent is a much broader one than that of merely obtaining a measure of his intelligence. A study of moral and character qualities has been urged by Healy, Bronner, Kohs and others. Healy

(15, 17) has been very insistent upon a thorough study of the whole personality of the delinquent in which the intelligence rating is only one item. The individual case studies of Healy and Bronner (22) are models of excellence for this type of approach. He has further shown the importance of mental analysis and indicated the frequency of mental conflicts as causative factors. The work of Goddard (21) at the Ohio Bureau of Juvenile Research has stressed greatly the importance of other mental abnormalities in addition to intelligence defect. He believes that a great number of delinquents are psychopathic, while at the same time they may or may not be of inferior intelligence. The use of the more recent character and personality tests will probably throw more light upon the other non-intellectual traits of delinquents. Investigators (Bridges, 26 and 27) are trying out such instruments as the Woodworth Personal Data Sheet, the Pressey X-O Tests and Kohs Ethical Discrimination Test.

These approaches to the study of delinquency are of unquestionable value, and the amount of space we have given them in this chapter is not to be interpreted as our estimate of its importance. The main topic of our whole book is the measurement of intelligence so that we merely mention these other approaches to the study of delinquency in order to remind the reader that the question of the intelligence of the delinquent is by no means the only one of importance.

The complicated factors at work in the causation of juvenile delinquency are nowhere better illustrated than in the work of Burt (25). His study seems to be the only one where a delinquent and a control group have been equated and compared. We cannot do better than quote from his conclusions: * "The following proves to be the order of importance of the various conditions we have reviewed: (1) defective discipline; (2) specific instincts; (3) general emotional instability; (4) morbid emotional conditions, mild

* From Burt, C. *The Young Delinquent.* By permission of D. Appleton and Company, publishers.

rather than grave, generating or generated by so-called complexes; (5) a family history of vice or crime; (6) intellectual disabilities, such as backwardness or dullness; (7) detrimental interests, such as a passion for adventure, for the cinema, or for some particular person, together with a lack of any uplifting pursuits; (8) developmental conditions, such as adolescence or precocity in growth; (9) a family history of intellectual weakness; (10) defective family relationships —the absence of a father, the presence of a stepmother; (11) influences operating outside the home—as bad street companions and lack or excess of facilities for amusement; (12) a family history of temperamental disorder—of insanity or the like; (13) a family history of physical weakness; (14) poverty and its concomitants; and, last of all, (15) physical infirmity or weakness in the child himself.

"Heredity appears to operate, not directly through the transmission of a criminal disposition as such, but rather indirectly, through such constitutional conditions as a dull or defective intelligence, an excitable and unbalanced temperament, or an over-development of some single primitive instinct. Of environmental conditions, those obtaining outside the home are far less important than those obtaining within it; and within it, material conditions, such as poverty, are far less important than moral conditions, such as ill discipline, vice, and most of all, the child's relations with his parents. Physical defects have barely half the weight of psychological and environmental. Psychological factors, whether due to heredity or to environment, are supreme both in number and strength over all the rest. Intellectual conditions are more serious than bodily, and emotional than intellectual; while psycho-analytic complexes everywhere provide a ready mechanism for the direction of overpowering instincts and of compressed emotional energy into open acts of crime."

Conclusions.—The problem of the intelligence of the delinquent is obviously one that has undergone marked changes during the last two decades in which it has been vigorously

attacked by means of intelligence tests. From a belief in a very large percentage of defective mentality among delinquents, we have come to suspect that this percentage is only about ten or fifteen, with some workers intimating that it may not differ from the percentage of feeblemindedness found in the general population. Nevertheless, most workers feel that the one most common factor associated with delinquency in general is defective mentality, and that, therefore, practical effort to understand, educate and segregate the feebleminded is of great importance in the solution of the problem of delinquency, even although such efforts cannot be expected to diminish radically the number of delinquents with which society has to deal.

As the emphasis upon defective mentality has decreased, there seems to be arising an attempt to find the difference between the delinquent and non-delinquent in other mental factors. Psychopathic disturbances, character and moral defects are hinted at as being of great importance to our problem, and in line with this thought we note the attempt to measure such factors objectively.

As to the importance of environmental factors, there seems to be the usual difference of opinion, but the general belief of the psychologist seems to be in the greater potency of original nature and he is inclined to see in many environmental features the results of original nature in the first instance. Poverty, alcoholism, lack of education, and the like are all to some extent the results of defective mentality and cannot be considered the sole or primary causes of anti-social conduct.

All reports so far have agreed in finding a much smaller percentage of delinquents of superior intelligence as compared with the amount of superior intelligence supposed to exist in the population at large. Also, there seems to be agreement in the finding of a larger proportion of defective delinquents among female as opposed to male delinquents. The type of crime a delinquent commits has evidently some relation to

his intelligence, but it is difficult to disentangle this one factor from all the others that are at work leading to the commitment of different sorts of crime. And, finally, the future success of a delinquent in the world is evidently conditioned by so many factors, other than intelligence, that we have not found, up to the present time, that the intelligence rating is at all prognostic of future success.

All these conclusions impress upon us the fact that the problem of delinquency is a very intricate one, and a very broad one. An individual may become delinquent through one of a great many causes. Other things being equal, however, an individual of defective mentality is more likely to become delinquent than one of normal or superior mentality.

REFERENCES

Aden, M. E. (26). "Preliminary Report of a Study of the Motor Ability of Delinquent Boys and Girls of the Los Angeles Juvenile Court." *J. of Delinq.*, 10, 351-367.

Adler, H. M. (28). *Tenth Annual Report of the Criminologist.* Dept. of Public Welfare, State of Illinois.

Adler, H. M. (29). *Eleventh Annual Report of the Criminologist.* Dept. of Public Welfare, Illinois.

Anderson, J. E. (21). "A Mental Survey of the Connecticut Industrial School for Girls." *J. of Delinq.*, 6, 271-282.

Asher, E. J. (27). "The Training Needs of Reform School Boys Experimentally Determined." *J. of Delinq.*, 11, 151-158.

Blanchard, P. and Paynter, R. H. (24). "The Problem Child." *Ment. Hyg.* 8, 26-54.

Bowler, A. C. (17). "A Study of Seventy-five Delinquent Girls." *J. of Delinq.*, 2, 156-168.

Bridges, J. W. (27). "A Study of a Group of Delinquent Girls." *Ped. Sem.*, 34, 187-204.

Bridges, J. W. and Bridges, K. M. B. (26). "A Psychological Study of Juvenile Delinquency by Group Methods." *Genetic Psych. Mono.*, 1, 407-506.

Bridgman, O. (13). "Mental Deficiency and Delinquency." *J. of Am. Med. Ass.*, 61, 471-472.

Bronner, A. F. (14). "A Research on the Proportion of Mental Defectives among Delinquents." *J. of Crim. Law and Crim.*, Nov.

Burt, C. (25). *The Young Delinquent.* New York.

Clark, W. W. (20). "Success Record of Delinquent Boys in Relation to Intelligence." *J. of Delinq.*, 5, 174-182.
Crane, H. W. (15). *Report on Feeblemindedness, Epilepsy and Insanity in Michigan.* Lansing, Mich.
Doll, E. A. (No date). *Psychologists' Report, N. J. State Prison.*
Doll, E. A. (21). "Mental Types, Truancy and Delinquency." *Sch. and Soc.*, 14, Nov. 26.
Dougherty, F. D. (26). "A Study of the Mechanical Ability of Delinquent Children of the Los Angeles Juvenile Court, 1925." *J. of Delinq.*, 10, 293-311.
Faber, H. K. and Ritter, A. J. (17). "A Mental and Physical Survey of a Group of Juvenile Delinquents." *Am. J. of Diseases of Children*, 14, 444-462.
Fernald, G. M. (16). "Report of the Psychological Work in the California School for Girls." *J. of Delinq.*, 1, 22-32.
Fernald, M. R. (17). "Practical Application of Psychology to the Problems of a Clearing House." *J. of Crim. Law and Crim.*, 7, 722-731.
Fernald, M. R., Hayes, M. H. and Dawley, A. (20). *A Study of Women Delinquents in N. Y. State.* New York.
Gifford, E. G. and Goddard, H. H. (12). "Defective Children in the Juvenile Court." *Training School Bull.*, 8, No. 9, January.
Gilliland, A. R. (16-17). "The Mental Ability of One Hundred Inmates of the Columbus [O.] Workhouse." *J. of Crim. Law and Crim.*, 7, 857-866.
Goddard, H. H. (15). *The Criminal Imbecile.* New York.
Goddard, H. H. (21). *The Bureau of Juvenile Research.* Ohio Board of Administration, Columbus, O.
Goring, C. (13). *The English Convict.* H. M. Stationery Office, London.
Haines, T. H. (15). "Point Scale Ratings of Delinquent Boys and Girls." *Psych. Rev.*, 22, 104-109.
Haines, T. H. (16). "Relative Values of Point Scale and Year Scale Measurements on One Thousand Minor Delinquents." *J. of Exp. Psych.*, 1, 51-82.
Haines, T. H. (16-17). "Feeblemindedness Among Adult Delinquents." *J. of Crim. Law and Crim.*, 7, 702-721.
Hall, G. E. (17). *Mental Examinations.* Eugenics and Social Welfare Bulletin, No. XI. State Board of Charities, N. Y.
Hamill, G. (23). "The Application of the Pintner Group Test to Misdemeanants." *J. of Delinq.*, 8, 158-168.
Healy, W. (15). *The Individual Delinquent.* Boston.
Healy, W. (17). *Mental Conflicts and Misconduct.* Boston.
Healy, W. (22). *The Practical Value of Scientific Study of Juvenile Delinquents.* Publication 96. Children's Bureau, U. S. Dept. of Labor.
Healy, W. and Bronner, A. (16). "Youthful Offenders." *Am. J. of Soc.*, 22, 38-52.

Healy, W. and Bronner, A. (22-23). *Case Studies.* Judge Baker Foundation, Boston.

Healy, W. and Bronner, A. F. (26). *Delinquents and Criminals.* Their Making and Unmaking. New York.

Healy, W. et al. (29). *Reconstructing Behavior in Youth.* New York.

Henry, M. B. (21). "Near Delinquents in the Public Schools." *J. of Delinq.,* 6, 529-548.

Hickman, H. B. (15). "Delinquent and Criminal Boys Tested by the Binet Scale." *Training School Bull.,* 2, 159-164.

Hill, H. F. and Goddard, H. H. (11). "Delinquent Girls Tested by the Binet Scale." *Training School Bull.,* 8, 50-56.

Kelley, T. L. (17). "Mental Aspects of Delinquency." *University of Texas Bull.,* No. 1713.

Kohs, S. C. (15). *The Practicability of the Binet Scale.* Publications of the Research Dept., Chicago House of Correction.

Kuhlmann, F. (26). *Report of Director.* Division of Research, State Board of Control, Minnesota.

Maris, C. L. (26). "A Psychological Survey of the Manitoba Juvenile Delinquent." *J. of Delinq.,* 10, 415-427.

Massachusetts (14). Report of the Commission for the Investigation of the White Slave Traffic, So Called. House 2281.

Mateer, F. (21). *Department of Clinical Psychology.* Bureau of Juvenile Research, Ohio. Publication No. 19.

Mathews, J. (23). "A Survey of 341 Delinquent Girls in California." *J. of Delinq.,* 8, 196-231.

McCaulley, S. (25-28). "One Hundred Non-Conformed Boys." *Psych. Clin.,* 16, 141-166.

McCord, C. P. (15-16). "One Hundred Female Offenders." *Training School Bull.,* 12, 59-67.

Memoirs, (21). The National Academy of Science, Vol. XV, Part III, Chap. 12.

Merrill, M. A. (26). "Mental Differences Among Juvenile Delinquents." *J. of Delinq.,* 10, 312-323.

Miner, J. B. (18). *Deficiency and Delinquency.* Baltimore.

Murchison, C. (26). *Criminal Intelligence.* Clark Univ., Worcester.

Ordahl, G. (16). "A Study of 341 Delinquent Boys." *J. of Delinq.,* 1, 72-86.

Ordahl, G. (17). "Mental Defectives and the Juvenile Court." *J. of Delinq.,* 2, 1-13.

Ordahl, L. E., and Ordahl, G. (17). "A Study of 49 Female Convicts." *J. of Delinq.,* 2, 331-351.

Ordahl, L. E. and Ordahl, G. (18). "A Study of Delinquent and Dependent Girls." *J. of Delinq.,* 3, 41-73.

Oseretzky, N. I. (28). "Die Minderjährigen Rechtsbrecher." *Zeitsch. f. Kinderforsch.,* 34, 53-74.

Otis, M. (13-14). "The Binet Tests Applied to Delinquent Girls." *Psych. Clin.,* 7, 127-134.

Pintner, R. (14). "One Hundred Juvenile Delinquents Tested by the Binet Scale." *Ped. Sem.,* 21, 523-531.

Pintner, R. and Toops, H. A. (17). "A Mental Survey of the Population of a Workhouse." *J. of Delinq.*, 2, 278-286.

Pintner, R. and Reamer, J. C. (18). "Mental Ability and Future Success of Delinquent Girls." *J. of Delinq.*, 3, 74-79.

Pyle, W. H. (14). "A Study of Delinquent Girls." *Psych. Clin.* 8, 143-148.

Richmond, W. (21). "An Industrial Institute Survey." *J. of Delinq.*, 6, 473-486.

Root, W. T. (no date). *A Psychological and Educational Survey of 1916 Prisoners in the Western Penitentiary of Pennsylvania.* Published by Board of Trustees of the Western Penitentiary.

Rossy, C. S. (15). "First Note on a Psychological Study of the Criminals at the Mass. State Prison." *Bulletin No. 13,* the Mass. State Board of Insanity, pp. 12-19.

Rowland, E. (13). "Report of Experiments at the State Reformatory for Women at Bedford, N. Y." *Psych. Rev.*, 20, 245-249.

Slawson, J. (26). *The Delinquent Boy.* Boston.

Spaulding, E. R. (14-15). "The Results of Mental and Physical Examinations of Four Hundred Women Offenders." *J. of Crim. Law and Crim.*, 5, 704-717.

Sullivan, E. B. (27). "Age, Intelligence and Educational Achievement of Boys Entering Whittier State School." *J. of Delinq.*, 11, 23-38.

Terman, L. M. and Knollin, H. E. (18). *Surveys in Mental Deviation.* California State Board of Charities and Corrections.

Weidensall, J. (16). *The Mentality of the Criminal Woman.* Baltimore.

Williams, J. H. (15). *A Study of 150 Delinquent Boys.* Research Lab., Buckel Foundation, Stanford Univ. Bulletin No. 1.

Williams, J. H. (15). *Defective, Delinquent and Dependent Boys.* Dept. of Research, Whittier State School, Bulletin No. 1.

Williams, J. H. (16). "Delinquent Boys of Superior Intelligence." *J. of Delinq.*, 1, 33-52.

Williams, J. H. (19). "The Intelligence of the Delinquent Boy." *J. of Delinq.*, Mono. No. 1.

CHAPTER XVII

THE DEPENDENT

Dependency is a very broad term and is often used to cover all those cases in which families or individuals are not self-supporting. Used in this broad manner, it would include not only the ordinary cases dependent upon public or private charitable agencies, but also those in institutions of various kinds such as feebleminded institutions, hospitals of various kinds, schools for the deaf, blind, cripples, and the like, and perhaps prisons and reformatories. In a narrower sense, however, the term is frequently used for individuals who are helped or supported by charitable agencies, such as children's homes, orphan asylums, almshouses, county infirmaries, and local charity organizations, to the exclusion of those afflicted with special disabilities, such as blindness, deafness, tuberculosis, and the like, as well as those convicted of delinquency. It is in this narrower sense that we shall use the term "dependent" in this chapter.

Almost all the intelligence testing of dependents has been confined to dependent children. We have very little direct knowledge of the mentality of the dependent adult. Most of the children tested have been those cared for by orphan asylums or children's homes. It is worth remembering in passing that the term "orphan" as applied to such children is very misleading. "Orphan" means strictly a child, both of whose parents are dead. As a matter of fact only a small percentage of the cases in these homes are orphans in the strict meaning of the term. A great many have one parent living and a still greater number have both parents living. The presence of the children in a home indicates, therefore, a

breakdown on the part of the family either because of some unfortunate accident to the breadwinner or else because of the incapacity of the breadwinner to support himself and family. The very fact that so many of the children have both parents living would lead one to suspect that in some cases the trouble is simply intelligence defect on the part of one or both parents, leading to an inability to compete with others in the ordinary course of life. This might lead us to suspect, therefore, a greater amount of feeblemindedness among such children than would be found in unselected school children.

Table 18 gives a summary of most of the studies available in which mental tests have been used. The first twelve studies are all concerned with children in so-called orphan asylums or county homes. The next two deal with dependent children referred to clinics by charitable agencies, and therefore represent more selected groups. Under dependent adults we have two studies of county infirmaries and one of adults aided by the associated charities.

Considering the first twelve reports of children, we note that the percentage of feeblemindedness varies from six to thirty-nine per cent with a median of about twelve per cent. This median of twelve is much lower than the median percentage of thirty for delinquents which we noted in the previous chapter. The percentage of seventeen reported by Haines (19) may be partially due to the fact that Mississippi has no special institution for the care of the feebleminded. The high percentage of thirty-four reported by Mateer (21) is in decided conflict with the other reports, and must be due to some radically different method of diagnosing feeblemindedness. Crane's (22-24) high percentage of 39 is based on only 33 cases in a children's refuge. In general, therefore, we may conclude that the amount of feeblemindedness among dependent children is less than that found among delinquent children, but still much greater than that generally assumed to exist among unselected school children. Furthermore, if

TABLE 18

DEPENDENT CHILDREN

Author	*No. Examined*	*F.M.*	*Back-ward*	*Normal*	*Above Normal*	*Institution*
Stenquist and others (15) ..	256	18.5	62.0	19.0	0.5	County Homes
Pintner (17)	106	5.7	46.2	34.9	13.2	A County Home
Hall (17)	2,142	6.7	...	...	...	23 Child-caring Institutions
Williams (18)	150	6.0	32.5	49.5	12.0	4 homes for children
Carlisle (18)	141	7.8	...	...	...	Orphan Asylum
Carlisle (18)	117	9.4	...	...	...	House of Good Shepherd
Terman and Wagner (18)..	68	6.0	29.0	53.0	12.0	Orphan Asylum
Haines (19)	270	17.0	...	...	...	Orphanages
Mateer (21)	1,603	33.7	3.9	13.8	0.6	34 County Homes *
Crane (22-24)	33	39.0	39.0	22.0	...	Refuge
Gesell (23)	198	18.0	21.0	52.0	9.0	County Homes
Davis (28)	1,051	15.0	49.0	31.0	5.0	Orphanages
Pintner (17)	82	19.5	39.0	36.6	...	Selected Clinic Group †
Bridgman (18)	133	26.0	40.0	34.0	...	Selected Clinic Group

DEPENDENT ADULTS

Crane (15) 3,334 cases in 79 county infirmaries interviewed and 21% were estimated to be feebleminded.

Brigger (16) 25 repeaters at the Associated Charities tested and 24% diagnosed as feebleminded.

Haines (19) 385 inmates of county poor farms examined and 36.6% diagnosed as feebleminded.

* In addition there are 46.8% of the cases diagnosed as potentially feebleminded or else "deferred diagnosis."

† In addition 4.9% diagnosed as doubtful.

the more recent suggestions that we have noted in the previous chapter with reference to the amount of feeblemindedness among delinquents were to be applied in the same way to the dependents, we would find in all probability that the percentage of feeblemindedness would be still further reduced.

When we examine the percentage of dependent children diagnosed as "backward," we note that, wherever figures are given, there is a very large percentage in this group. Poor mentality, even although it may not amount to actual feeblemindedness, would seem, therefore, to be a characteristic of this type of child. Again we find a fair proportion of children rated "above normal." Although our data here are not very adequate, it is interesting to notice that such cases were very seldom reported for delinquent children. It would seem, therefore, that our chances of finding superior children among dependents are much greater than among delinquents. And this would be reasonable, because there are many accidental causes which may lead to dependency in children, who come from a stock possessing good intelligence traits.

The report of Davis (28) deals with 1,051 cases in Texas orphanages and compares these with 504 public school children in the same region. All these cases were given two group tests, the Dearborn and the Haggerty. The striking difference in mentality between the dependent and non-dependent children can be seen from the percentage distribution of I.Q.s:

	Dearborn Test		*Haggerty Test*	
I.Q.	*Orphanage*	*Public School*	*Orphanage*	*Public School*
Above 139	0	0.4	0.1	1.6
120-139	1.0	7.1	3.4	10.9
110-119	4.0	13.1	6.8	12.3
90-109	31.0	47.6	30.2	43.7
80-89	27.0	18.4	23.1	18.0
70-79	22.0	9.0	19.0	8.9
Below 70	15.0	4.4	17.4	4.6

About 60 per cent of the orphan children fall below an I.Q. of 90, whereas only half that percentage of the public school cases fall below this point.

The two reports by Pintner (17) and Bridgman (18) of special cases sent to a clinic for examination show a much larger percentage of feeblemindedness than is found in orphan asylums in general. They are obviously cases that have been referred for examination because of suspected defect or because they were problem cases in some form or other.

Blanchard and Paynter (27) report the I.Q.s of 80 children belonging to 23 families in receipt of charity in a large city. The children in this case are not in an orphanage but are at home. The percentage distribution of I.Q.s is:

Below 70	*70-80*	*80-90*	*90-100*	*100-109*	*110-120*	*Above 120*
11.2	10.0	31.3	30.0	11.6	6.3	0

Here again we see the same piling up of poor intelligence.

The reports of dependent adults show a larger proportion of feebleminded than the children's reports, but from this we can draw no conclusion as to the relative amount of mental defect among dependent adults and children. The cases reported by Crane (15) are infirmary cases and represent the most extreme cases of dependency. Furthermore, the study was made relatively early, and later and more conservative methods of diagnosis might lead to a reduction of the percentage for almshouses in general. The percentage of feebleminded found by Haines (19) in the poor-farms of Mississippi is very large, and a partial explanation of this is that Mississippi at the time of the survey made no special provision for feebleminded individuals. The other report by Brigger (16) represents cases of more or less chronic dependency helped by a charitable organization, and this represents a very selected group of dependent adults. We would expect to find the incidence of feeblemindedness rather high in such a group, but again, later and more conservative methods of diagnosis

might decrease the amount. We have no study, so far as the writer is aware, of the intelligence of the general run of dependent adults, such as are dependent upon the more or less regular assistance of our local charity organizations. We can only infer from the studies of the difference in intelligence between children of different social groups, and from the results for dependent children, that the percentage of defectives would be somewhat larger for the dependent adult than is to be found in the general population.

Dependency and delinquency are closely connected. The overlapping between the two groups is great. One common factor would seem to be a somewhat defective mentality. Poverty and neglect are associated with defective mentality. The combination of poverty and lack of intelligence is frequently met with in delinquency. It would, therefore, seem to be wise to examine thoroughly the intelligence of dependent children and make adequate provision for the mentally deficient in order, as far as possible, to forestall the delinquency that will probably result. And, further, because there seems to be a certain percentage of superior children among the dependent cases, it would be wise and just to make adequate provision for such cases, so that they might have opportunities for education and advancement in proportion to their intellectual capacity.

REFERENCES

Blanchard, P. and Paynter, R. H. (27). "Socio-Psychological Status of Children from Marginal Families." *The Family*, March, 3-10.

Bridgman, O. (18). *An Experimental Study of Abnormal Children, with Special Reference to the Problems of Dependency and Delinquency*. University of California Publications in Psychology, Vol. III, No. 1.

Brigger, G. (16). "A Study of Twenty-five Repeaters at the Associated Charities, Portland, Oregon." *J. of Delinq.*, 1, 187-194.

Carlisle, C. L. (18). "The Causes of Dependency, Based on a Survey of Oneida County." *Eugenics and Social Welfare Bulletin*, No. 15. State Board of Charities, N. Y.

Cobb, M. E. (22). "The Mentality of Dependent Children." *J. of Delinq.*, 7, 132-140.

Crane, H. W. (15). *Report on Feeblemindedness, Epilepsy and Insanity in Michigan.* Lansing, Mich.

Crane, H. W. (22-24). *Report of Bureau of Mental Health and Hygiene.* Biennial Report of the North Carolina State Board of Charities and Public Welfare.

Davis, R. A. (28). "The Study of Orphan Children in Texas." *J. of Delinq.,* 12, 1-21.

Gesell, A. (23). *The Pre-School Child.* Boston.

Haines, T. H. (19). *Mississippi Mental Deficiency Survey,* Miss. Mental Hygiene Commission.

Hall, G. E. (17). "Mental Examinations." *Eugenics and Social Welfare Bulletin,* No. 11. State Board of Charities, N. Y.

Mateer, F. (21). *The Bureau of Juvenile Research,* Ohio Board of Administration, Publication No. 19, pp. 5-50.

Pintner, R. (17). "The Mentality of the Dependent Child." *J. of Ed. Psych.,* 8, 220-238.

Stenquist, J. L. et al. (15). "The Intellectual Status of Children who are Public Charges." *Arch. of Psych.,* No. 33.

Terman, L. M. and Wagner, D. (18). "Intelligence Quotients of 68 Children in a California Orphanage." *J. of Delinq.,* 3, 115-121.

Williams, J. H. (18). *The Intelligence of Orphan Children.* Surveys in Mental Deviation. State Board of Charities and Corrections, California, 46-81.

CHAPTER XVIII

THE DEAF

The Extent of the Problem.—The handicap of deafness is such as to make a deaf person, particularly if uneducated, very different from a normal hearing individual. Deafness isolates him very much from the world in general, because he lacks the easy and quick method of communication between man and man afforded by speech. The deaf have, thus, always been looked upon as peculiar. Their general lack of speech is called "dumbness" and in American slang the words "dumb" and "dummy" have acquired meanings which have obviously arisen because of the difficulty which the deaf experience in understanding the ordinary speech of their fellowmen. Again, the deaf are relatively few in number. This also has led to the general misunderstanding of them which exists among hearing individuals, and has intensified their isolation. If their numbers were much greater, the chances for the hearing individual to meet and learn to understand them would be more numerous. The 1920 Census (23) gives the total number of "deaf and dumb population in the United States" as 44,885, or about .05 per cent of the total population. This makes the problem of the deaf in regard to numbers at least a relatively minor one and explains why they are for the most part misunderstood by the population at large. Best's (14) analysis of the problem leads him to believe that the number of deaf in this country is decreasing, although very gradually.

Of the 44 thousand deaf listed by the Census in 1910, nineteen thousand replied to questionnaires, showing 39.3 per cent in which deafness was congenital (i.e., existing at birth) and

60.7 per cent acquired after birth. The percentage of acquired deafness is undoubtedly exaggerated because of the difficulty of determining whether a child can hear shortly after birth and the general tendency to believe that deafness begins when the family first notices something peculiar about the child with reference to his inability to speak after the first twelve months or so. An analysis of seventeen thousand cases shows about 24 per cent reporting speech used as a means of communication, and 72 per cent not using speech as a means of communication. The number of deaf children in schools is given by Hall (21) as 13,779, the greater proportion or 80 per cent of which are in public residential schools.

Study of the Deaf.—The deaf have for a long time been studied by scientists and educators. For the most part, however, this study has been concentrated upon two types of problems, namely, the medical and the instructional. Under the medical we include investigations as to the causes and possible cure of deafness, and under the instructional the methods of teaching the deaf. Very little psychological work has been done and only within recent years have psychological tests been given.

The history of the education of the deaf is fascinating and in some respects dramatic. Best (14) gives a good account of it and Jones (17) outlines the work in America during the past one hundred years from the time of the opening of the first public residential school for the deaf in 1817, while Fay (93) gives a detailed history of all the American schools. The absorbing interest of the educators of the deaf during the past century has been the question of methods of instruction, and more particularly what means of communication between themselves and their hearing fellows the deaf should be taught to use. At times this question of means of communication has usurped the stage so completely as to crowd out all the other important problems. Indeed, the outside observer, at all inclined to be facetious, might well remark that the educators of the deaf did not seem to care what the deaf learned

but were very much concerned how they learned it, whether by oral or manual methods. The fight between the oralists, who advocate the teaching of speech and lip-reading, and the manualists, who advocate signs and finger-spelling, has absorbed much of the energy of the teachers of the deaf and, perhaps, diverted their attention from other equally important matters. At the present time a great majority of the deaf are taught some speech and lip reading, whether supplemented or not by a knowledge of signs and finger-spelling.

Earlier Psychological Studies.—A detailed account of the earlier psychological studies is given by Reamer (21) and by Pintner and Paterson (17). Here we may simply note that they were concerned with such miscellaneous topics as brain defects and the so-called speech center, the need for classification of deaf children, the problem of the feebleminded deaf child, the peculiar deaf child, the spelling ability of the deaf, anthropometric measurements, memory, drawing ability and the like. Few of the studies were extensive or even scientific. The work of Love (96 and 12) is mainly medical in character, but he speculates upon the mental characteristics of deaf children. The head measurements are smaller than the hearing—"due to neglect during the period of language formation"! He says, "absence or great defect of hearing arrests mental progress. . . . The physical counterpart of this intellectual basis is that the deaf child at 7 years has a smaller brain than the hearing child of the same age," and again, "the absence of hearing, on which thought processes so much depend, finds no adequate compensation by increased activity of the other senses." All of this is very speculative and very questionable, particularly without adequate psychological tests.

The nearest approach to anything resembling intelligence tests seems to be the tests described by Greenberger (89) for finding out whether the child entering an institution was of fair mentality or not. He suggests that the child be shown attractive picture books and that the examiner watch what

the child does. If the child remains perfectly apathetic, it is a bad sign; but if he brightens up at the sight of the books and maintains an interest in them for a period of time, it is an indication of fair mentality that can be improved by training. This is in no sense a test in the modern meaning of the term, but the idea was there, and unfortunately was not further developed. The application of psychological tests proper had to wait for the coming of the psychologist. Macmillan and Bruner (06) seem to have been the first psychologists to apply mental tests. They used the well-known cancellation of "A's" test, perception of size by sense of touch, sensitivity for lifted weights, a few memory tests as well as physical and sensory-motor tests. In the tests of so-called higher mental functions, they found the deaf to be generally inferior. They suggest that, "this inferiority of the deaf on the mental side perhaps means no more than that the child is from three to four years less mature than the hearing child of his age, and that his date of maturity will be correspondingly delayed." This conclusion is very much open to question and in the light of the results to be mentioned later seems very doubtful. In all probability the average deaf adult remains somewhat inferior to the average hearing individual, regardless of age.

Pintner and Paterson (15) seem to have been the first to apply well-recognized intelligence tests or scales to deaf children. They applied the Goddard Revision of the Binet Scale to 22 deaf children. Various methods of communicating with the children were tried, namely, writing, speech, manual spelling and signs or any mixture of these. As a result of this experiment, they concluded that this type of intelligence scale was totally unsuited for the testing of a deaf child. The results showed the deaf child on an average to be four and one-half years retarded. The inadequacy of language tests as tests of intelligence for the deaf is, however, the vital point that is brought out by the study, and it

was this work with the deaf that stimulated them to construct their Performance Scale.

Comparison of Intelligence of Deaf and Hearing.—When we make such a comparison we must be careful to exclude all tests in which language is involved, whether in the directions to the subject or in the test material itself. Language in the hearing child we regard as a medium through which we may probe into his mental make-up. Although opportunities for acquiring language may vary greatly from individual to individual among hearing children, yet there is enough in common to make it a basis for intelligence testing. Opportunity for hearing language, and later on for reading, is so widespread that slowness or disability in language is itself indicative of mental inferiority. Not so with the deaf child. For him, the English language is a subject of instruction in the school and his progress in it is painfully slow. A test involving language immediately becomes for the deaf a subject-matter test depending very much upon the amount and length of schooling, the effectiveness of the schooling, and the like. Of course, it is indirectly, like all subject-matter tests, also a test of intelligence and can be so used whenever we have children of equal amounts of schooling, as is the case with educational tests among the hearing. When, however, we want to measure the intelligence of deaf children of different ages and with different amounts of educational opportunity, we must go back of our language tests to something more common to deaf children in general. Hence the necessity for using non-language and performance tests.

Pintner and Paterson (15 and 16) and Reamer (21) compare the deaf and hearing on non-language group tests. The deaf fall decidedly below the norms for the hearing in all cases. With about 1,000 cases on the Digit-Symbol and Symbol-Digit Tests, the percentage of deaf boys reaching or exceeding Pyle's median for hearing boys is 24 and 31 per cent respectively; for deaf girls as compared with hearing girls 10 per cent for both tests. Pintner and Paterson conclude that

on the whole the deaf child is about 3 years behind the hearing on these two tests. This amount of retardation need not be found in every school. This is shown by Newlee (19) who repeated the tests mentioned above with 85 deaf children in a day school and found them about up to the norms for the hearing. We should expect to find differences in intelligence among groups of deaf children just as we find them among groups of hearing children. The Pintner-Paterson group of 1,000 cases is a much more representative sampling of deaf children in general than the small select group of Newlee.

Reamer (21) used a much more comprehensive non-language intelligence test. At all ages, where representative groups of deaf and hearing children could be compared, the deaf are much below the hearing. The conservative conclusion is that the deaf are about two years behind the hearing and, if the younger children are omitted as being more highly selected among the deaf, we have a retardation of two and a half years. In such a special ability as visual memory for digits, the deaf are greatly inferior to the hearing (Pintner and Paterson 17), showing no compensation in this ability for their lack of hearing.

A nation-wide survey of deaf children (Day, Fusfeld and Pintner, 28, and Pintner, 28) by means of the Pintner Non-Language Test gives us the most comprehensive comparison of the deaf and hearing. All children, age 12 and above in 41 schools were given this test, making in all a total of 4,432 cases. The distribution of these cases according to the nine geographic divisions of the United States resembles the distribution of the total deaf population as given by the 1920 census. The number of cases is large and probably amounts to as much as 50 per cent of all deaf children of those ages. In all probability this sampling of deaf children is very representative of deaf children of like age in general. Table 19 gives a comparison of the deaf and hearing at ages 12 to 15 inclusive. The sampling of hearing children, though large, is probably not as representative as that of the deaf. The

TABLE 19

COMPARISON OF DEAF AND HEARING ON THE PINTNER NON-LANGUAGE TEST

Age	*12*	*13*	*14*	*15*
Mean score—hearing	321	348	362	364
No. of cases—hearing ...	1,361	1,295	1,120	700
Mean score—deaf	257.7	275.6	300.5	320.1
No. of cases—deaf	547	608	678	590
Difference of means	63.3	72.4	61.5	43.9
S.D. difference	5.22	5.32	5.31	6.13
Ratio	12.12	13.59	11.59	7.16

cases are drawn largely from the elementary school and the fifteen-year-old group may not contain a sufficient number of cases of superior intelligence. At any rate, we may be certain that the deaf are not being compared with hearing children of superior intelligence. A study of the table shows us that at all four ages the hearing are decidedly superior to the deaf and that the differences between the means are statistically reliable. Another way of comparing the two groups is to interpret the means of the deaf in terms of the hearing norms. The mean score of the 12-year-old deaf is 258 and this is about the mean for the 10-year-old hearing and from this we may calculate an I.Q. of 83. Hence we have:

Deaf	*Hearing Equivalent*	*I.Q. of Deaf*
Age 12	10	83
Age 13	10-6	81
Age 14	11-0	79
Age 15	12-0	80

From this comparison we note that the deaf are from two to three years retarded, and this is best represented by an I.Q. of about 80.

All of the studies so far mentioned have found the deaf

very decidedly below the hearing on all kinds of intelligence tests, but in Scotland, Drever (29) and Drever and Collins (28) find the deaf up to or even a little above the hearing norms on their Scale of Performance Tests. These findings are based on tests of 1,474 deaf children between the ages of 5 to 16. Only a brief report of this work has so far appeared and it is impossible to explain the discrepancy between these results and the great number of studies made in the United States.

Mentality of Acquired and Congenital.—The same authors quoted above (Reamer, Pintner and Paterson), give comparisons between those children who are born deaf and those who become deaf after birth. Taken as groups, they show no difference in their average general intelligence. Pintner (28) gives the following median indices of intelligence for 2,423 deaf children, ages 12 to 15, according to the age of onset of deafness:

Age	*Unknown*	*Birth*	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*Total*
Intelligence index.	48.4	50.0	49.7	50.5	49.6	51.2	52.7	51.1	51.6	51.5	53.1	56.0	56.0	53.0	53.0	50.0	50.3
n	201	1,129	222	259	201	104	75	74	42	43	33	17	17	6	4	1	2,423

An inspection of these indices shows only a very slight tendency to rise after age 2, but the rise is so slight and so irregular that we cannot be certain that it represents any real increase in intelligence. We are bound to conclude that there is no difference between the intelligence of those born deaf and those who become deaf after birth. Nor is there any real correlation between age of onset of deafness and intelligence. Upshall's (29) two correlations between these two variables for two different groups are —.02 and +.09.

Orally and Manually Taught Pupils.—The comparison of the intelligence of deaf children taught by the oral method

with those taught by the manual method shows a superiority of the orally taught pupils. This is obviously due to the selective policy of the schools, by means of which the brighter children are generally chosen for oral work. In combined schools (i.e., schools which teach by both methods) all children are generally taught by the oral method at first and those who fail to succeed are relegated to the manual classes. Reamer's (21) median mental index for 1,753 oral pupils is 51 and for 211 manual pupils only 43. Pintner (28) reports median indices as follows:

Method of Instruction	*Intelligence*	*n*
Oral	52.0	1,845
Manual	41.2	186
Combined	46.9	392
Manual and combined	44.8	578
Total	50.4	2,423

From a detailed distribution of the cases he finds a selective policy at work, whereby three times as many non-oral children are found having low intelligence ratings as compared with oral children. And, conversely, more oral than non-oral children are found with high intelligence ratings.

As to the value of the oral method as a practical means of communication for deaf children, none of these psychological studies is concerned directly with the problem. The only attempt in this direction seems to have been made by Binet and Simon (09) with conclusions very unfavorable to the oral method as of value in after life. It is interesting to note that they comment upon the difficulty which the oral method encounters because of the lack of intelligence of the deaf, a lack surmised by the authors but not measured. This lack of intelligence has since become obvious in standard intelligence tests.

Residential and Day Schools.—Most of the deaf children in the United States are educated at present in residential schools. A small percentage is educated in day schools. The establishment of the day school is relatively recent. Upshall (29) has made a very intensive study of children in these two types of schools. His results are based on tests of intelligence, educational achievement and audition, as well as data gathered by a questionnaire. Upshall paired 311 day school cases with 311 institution cases for C.A. and intelligence in order to measure the influence of the two types of schools. The main results of his intensive study can best be understood by quoting some of his conclusions:

"1. The difference between mental ability of Day School pupils and Institutional pupils is statistically reliable. The Day School selects the brighter children.

"2. There is a statistically reliable difference between the Day School pupils and the Institutional pupils in educational achievement as measured by the Pintner Educational Survey Test.

"3. Children who attend the Day Schools have, in general, a greater degree of residual hearing than children who attend the Institutions. Statistical certainty is attained.

"4. Superiority in degree of residual hearing is accompanied by a later age of becoming deaf and a greater number of years spent in a school for normal hearing pupils. Statistical certainty is attained.

"5. When the important factors of age and mental ability are made equal there is still a real difference in favor of the Day School pupils in educational achievement.

"6. When, in addition to the factors just named, only those children in the Day Schools who have been deaf from the age of one year and who have never attended a school for normal hearing children are compared with similar children in the Institutions, the difference be-

tween the two types of schools is reduced slightly over fifty per cent.

"7. The difference is still in favor of the Day Schools. With the 83 cases used there are 99 chances in 100 that the true difference would be above zero.

"8. Since the process of equating, as indicated in conclusion 6, also equalized the factors of residual hearing, age of starting to school, and years spent in a deaf school, the conclusion must be reached that the chances are very great that the Day Schools are superior to the Institutions in the type of education which is measured by the Pintner Educational Survey Test."

The Language Ability of the Deaf.—That language ability is no adequate measure of the intelligence of the deaf child is shown by the comparison of deaf and hearing in purely language tests. Pintner and Paterson's (16) comparison by means of the Trabue language scale shows that very few deaf children (only 6.4 per cent) score above fourth grade ability of hearing children. At every age and grade the difference is very great. The same authors (16) have also compared the deaf and hearing on the Woodworth and Wells Directions Tests and conclude that "the average deaf child's ability to comprehend the language involved in these tests is about equal to that of the average hearing child between the ages of six and eight." Again Pintner (18) shows how terribly slow is progress in language ability in the deaf in his attempt to measure gain in language ability over a period of six months, whereas such gain can be measured in hearing children. Reamer (21) shows how the general educational ability of the deaf as tested by standard educational tests lags behind the ability of the hearing. She finds an average retardation of five years in educational ability as compared with hearing children (contrast this with the retardation of two years in mental ability on non-verbal tests). The results of the nation-wide survey of the deaf (Pintner, 28;

Day, Fusfeld and Pintner, 28) show a similar wide discrepancy between the educational scores of deaf and hearing. Turning the deaf mean scores into hearing age equivalents and calculating E.Q.s we have:

Deaf	*Hearing Equivalent*	*E.Q. of Deaf*
Age 12	7-9	65
Age 13	8-1	62
Age 14	8-9	63
Age 15	9-0	60

This gives us as a maximum an E.Q. of 65. If we compare this with the I.Q. of 80 on the intelligence test we have a measure of the relatively greater retardation in educational achievement as contrasted with intelligence. This shows very clearly the handicap of deafness in the acquisition of language. Whether the E.Q. of 65 could be brought up to the I.Q. of 80 by more efficient training, it is impossible to say.

McManaway (23) and Hall (29) also report the results of standard educational tests given to deaf pupils. McManaway finds his 13-year-old deaf equal to the norms for 8-year-old hearing, and Hall's Preparatory and Freshmen College groups are equal to hearing children in Grades IX and X. All the studies mentioned in this paragraph agree in finding the adventitiously deaf better in language tests than the congenitally deaf. The adventitiously deaf who lose their hearing after age five or six show decidedly superior language ability. Thus Pintner (28) gives the following median educational indices for different ages of becoming deaf:

Age	*Unknown*	*Birth*	*0*	*1*	*2*	*3*	*4*	*5*
Educational index.	45.5	48.8	47.4	48.1	50.0	51.9	56.3	61.2
n	201	1,129	222	259	201	104	75	74

6	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*Total*
63.6	62.6	63.3	71.0	68.0	68.0	67.0	70.0	49.8
42	43	33	17	17	6	4	1	2,423

There is a definite tendency here to increase in index with increase of age of onset of deafness. Every year of hearing after the second birthday seems to give the child a better chance for educational achievement in later life. With the hearing child speech begins at about twelve months, increasing rapidly thereafter. Each year of speaking and hearing gives the child a rapidly increasing vocabulary, such as the child born deaf cannot obtain. The advantage of thus acquiring some language before deafness sets in is evidently very great and is something that affects the language ability in later life. The most marked rise occurs between the third and fifth year. On the average a child who does not become deaf until after his fifth birthday will have a distinct advantage in language work in later life.

With reference to method of instruction, Reamer points out that, although manually taught pupils show markedly inferior educational ability in comparison with the orally taught, yet if we take into consideration their intelligence both groups are working equally well. The oral method does not seem to lead to a better comprehension of written English. "Oral and manual pupils of the same mental caliber do equally well on the educational test." Pintner (28) comes to the same conclusion from his analysis of the results of the large survey. He says, "Children not educated by the oral method alone seem to achieve as much in written language work as do orally educated children, when we take into consideration their intelligence. Obviously these findings challenge very definitely the assumed superiority of the exclusively oral method in deaf schools. If the oral method results in better language, then this should be measurable in written work. It has not been found so in the present case. This does not mean the abandonment of the oral method, because the manual method has not shown itself in any sense superior, but merely just as good. The claim for the continuance of oral work in deaf schools must, therefore, shift to the additional advantage that orally taught children have in

speech and lip-reading as compared with manually-taught children."

Motor Ability.—Long (31) conducted a study of the motor abilities of deaf children. He added tests of grip and balance to five of the six tests of the Stanford Motor Skills Unit and with this battery carried out measurements in matched groups of deaf and hearing children. Each group consisted of 87 subjects (51 boys and 36 girls) ranging in age from 8 to 18, and the deaf and hearing were equated in pairs for age, sex and race. Groups were compared on the basis of mean scores.

Over the entire battery the deaf boys averaged somewhat superior to the hearing, while the reverse was true for the girls. With the sexes combined, the resulting deaf and hearing groups would be about as nearly equivalent in average motor abilities as they well could be within the errors to be expected from random sampling. The only test in which both sexes of the deaf were superior were the Brown Spool Packing and the Motility Rotor, and this was offset by superiority of both sexes of the hearing in the Serial Discriminator and balancing. The balancing test was the only one in which the differences between the means of the groups were significant (difference between means more than three times the standard deviation of the difference).

The superiority of both groups of the hearing with the Serial Discriminator is rather suggestive. This is a reaction time test depending on the speed with which subjects respond to a series of stimuli. The stimuli used here were the digits 1, 2, 3 and 4 appearing in random order, and it is likely that skill in performance is linked up with intelligence and school achievement.

Practical Uses of Tests in Deaf Schools.—Little practical use of tests in deaf schools is reported, although suggestions are made which have to some extent been followed. Classification into homogeneous groups is stressed by Pintner (18), who shows in a particular case how a group of 29

young deaf children tested by means of an individual scale, The Pintner-Paterson Performance Scale, should be grouped into the three beginning classes which were available in the school in question.

The first mental-educational survey of a school was reported by Pintner and Paterson (16), to be followed by the later extensive survey of 26 deaf schools totalling 2,172 children undertaken by Reamer (21). The latter study shows the same great difference in intelligence among schools for the deaf as we have met with in schools for the hearing. The median mental indices of the schools ranged from 37 to 63. The percentage of children rating "very bright" ranges in the schools examined from 5.6 to zero; the percentage of cases rating "dull" from zero to 5.7 per cent. The percentage of cases rating "normal" varies from 71 per cent in the highest school to 22.2 per cent in the lowest. Reamer also makes use of combined mental-educational tests to point out the necessity for more efficient methods of instruction and administration.

Conclusions.—The author has summed up elsewhere (Pintner and Paterson, 18) his opinions based upon his psychological tests of the deaf, and the rest of this chapter is substantially a quotation from this earlier report.

A careful study of the results obtained from our series of experiments reveals at once two outstanding facts: First the startling deficiency of the deaf in their ability to comprehend and handle printed and written language; and, second, the general mental inferiority of the deaf as a group.

At an early stage of our work we discovered that this language deficiency was due almost wholly to the lack of normal social intercourse. With the hearing child incessant social intercourse leads to the development of a speaking vocabulary which serves as a basis for the rapid acquisition of reading ability. Due to the fact that the deaf child is cut off from such language environment he develops a vocabulary very slowly; as a consequence of this he does not have a medium

into which he can translate the ideas of the printed page. The difficulties encountered by the deaf in the acquisition of language (regardless of methods of instruction) seem to be well-nigh insurmountable. The fact that language ability is not a reliable index of the native mental capacity of the deaf child makes it necessary to keep the measures of mental ability—that is, our performance tests—separate from measures of language ability. In the case of the deaf child the latter are measures of acquired habits. These acquired habits are greatly affected by the vicissitudes of home and school life, and therefore, do not give us an index as to the child's native intelligence.

In general, the deficiency in language ability and in tests involving the utilization of auditory processes is a characteristic of both the congenitally and adventitiously deaf. Of the latter group, however, those who possessed hearing after the age of four or five are superior, on the average, to the rest of the deaf. This merely shows how necessary is the sense of hearing for the development of language ability.

In considering the mental inferiority of the deaf, the question naturally arises as to the cause of this inferiority in activities that presumably do not involve audition. The results indicate that the adventitious and congenital groups show the same general inferiority in these traits. However, it does not follow that this result is produced by a single cause. It is necessary, therefore, to consider the question for each group separately.

Reliable authorities state that 60 to 70 per cent of deafness occurs after birth. Now, the two chief causes of adventitious deafness are cerebral meningitis and scarlet fever. Since both these diseases are known to affect in many cases the mentality of the normal hearing child, we are tempted to conclude that the mental retardation of many of the adventitious group is caused by the disease which caused the deafness. Instead of deafness being the cause of mental inferiority, we find that the disease which produced the deaf-

ness caused at the same time the mental backwardness. We seem, therefore, to have a partial explanation for the inferior mentality of the adventitious group.

In regard to the congenital group, there is no specific disease which can be pointed to as a cause. It is possible that congenital deafness may be due in some cases to pathological prenatal (non-hereditary) causes, which at the same time affect the capacity for mental development. Two other possible explanations may be offered:

1. That loss of hearing may preclude normal mental growth, even in those traits which presumably do not depend upon hearing for their development.
2. That congenital deafness, on the whole, occurs more frequently in families of inferior mental ability.

The first possibility seems untenable, since there are a number of cases of deaf children showing ability equal to, and in some cases superior to, that of average hearing children. One might reply that these superior congenital deaf children might have shown even better mental ability if they had once possessed hearing. According to this point of view, the average ability of the congenital group would have been higher (even equalling the average ability of hearing children) had they all retained their hearing. In spite of this, however, it does not seem that mere absence of hearing itself is sufficient to explain deficiencies in activities which develop for the most part independently of auditory processes.

Therefore, it is necessary to inquire into the truth of the second possibility. At the outset of such inquiry, we find that there is little or no evidence concerning the family histories of our congenital cases. No investigations have been made which deal with the mentality of these families. Goddard has shown only a small amount of deafness occurring in the inferior families which he investigated. He is unable to draw any conclusions from his data in regard to the point we have raised. But he was investigating families in which

the principal defect was feeblemindedness. It is obvious that this is not the line of attack that must be made to solve the problem. We must start with congenitally deaf children and trace back their family histories and measure in some way the mentality of the people comprising the family group. Such an investigation is urgently needed, and until such is made we must wait for a satisfactory explanation of the inferior mental ability of the congenitally deaf.*

Although the reason for the mental retardation of the deaf is unknown, the fact that they are, as a group, mentally backward is quite evident. This has been the outstanding result of all the mental tests. The deaf child is, on the average, two or three years retarded in mental development as contrasted with the hearing child. This fact has been brought out both in the individual tests and in the group tests. Tests of space perception, of meaningful perception, of learning, and tests of attention have all shown the same results. These indicate a general inferiority in various mental traits, a general lowering of mental capacity rather than an inferiority in specific traits. Hence this inferiority does not seem to have been caused by the mere fact of deafness. Where mere lack of hearing affects the deaf child's performance (as in visual memory for digits and in tests of language ability), we find the deaf child much more seriously affected and handicapped.

This retardation of two or three years seems to exist all through the school career of the deaf child. It is not an initial retardation which is later on overcome. We find no evidence for a commonly accepted view that a deaf child starts out behind the hearing child and "catches up" later on. We do not believe that the average deaf child ever "catches up" to the average hearing individual. It must be borne in mind that we are speaking of the average and are well aware of the fact that there are exceptional deaf individuals who are distinctly above the average hearing individual.

* See the later work of Pintner and Osborn mentioned in Chapter XXIV of this book.

This mental retardation applies to the mental ability of the deaf child, as far as we are able to measure it at the present time. It does not apply to the language ability of the deaf child. We feel it imperative to keep language and mental tests entirely separate when dealing with deaf children. The measurement of language ability shows the deaf child to be about four or five years retarded. The enormous handicap under which the deaf child labors in the acquisition of language has been brought out clearly for the first time in such measurements as we have described. For the psychologist, the interest centers largely in the important part that audition is thus shown to play in the language development of the hearing. For the educator of the deaf, it is of interest in pointing out that progress in language must necessarily be slow; that great emphasis must be placed upon it, and that he must realize the limits in this direction and yet not be discouraged. Recognizing these limits, he will not fall into the fallacy of neglecting all else in the attempt to attain an impossible degree of perfection in language, but rather will he be satisfied with little and slow progress, and at the same time utilize all the other time of the deaf child in teaching him those things which will best enable him to take his place as a socially desirable citizen.

The science of education is being based more and more on the theory that it is the function of education to make each individual socially effective. Due to the rise of psychological measures of individual differences, educators are realizing that special kinds of education must be provided in order that each child may be developed to the limit of his capacity. In view of our results, it would seem that deaf education should more and more emphasize industrial training. Our results indicate that the deaf child, because of his language deficiency, is not very well suited for academic instruction. Much of the instruction in geography, formal grammar, physiology, history, Latin and algebra seems to be in large part a useless expenditure of energy. On the other hand, the deaf

child is more nearly on a footing with his hearing brother in those motor capacities that are fundamental for industrial success. Hence the deaf child has a greater chance of becoming socially effective if given a real opportunity for thorough and adequate industrial training. The academic instruction should be designed primarily to aid the deaf pupil in mastering the problems in his trade. English instruction should be designed to equip the child, so far as possible, for simple social and business intercourse.

REFERENCES

Best, H. (14). *The Deaf.* New York.

Binet, A. et Simon, T. (09). "Peut-on enseigner la parole aux sourds-muets?" *L'année psych.,* 15, 373-396.

Census (23). Deaf-Mutes in the United States: 1920. Bureau of the Census, Department of Commerce, Washington.

Day, H. E., Fusfeld, I. S. and Pintner, R. (28). *A Survey of American Schools for the Deaf.* National Research Council, Washington, D. C.

Drever, J. (29). "Intelligence Tests for the Deaf." *Teacher of the Deaf,* 27, 163-167.

Drever, J. and Collins, M. (28). *Performance Tests of Intelligence.* Oliver & Boyd, Edinburgh.

Fay, E. A. (93). *Histories of American Schools for the Deaf.* Three vols. Volta Bureau.

Greenberger, D. (89). "Doubtful Cases." *Am. Annals of the Deaf,* Vol. XXXIV, No. 2, p. 93 et seq.

Hall, P. (21). *Education of the Deaf.* Bulletin No. 14, Dept. of the Interior, Bureau of Education, Wash.

Hall, P. (29). "Results of Recent Tests at Gallaudet College." *Am. Annals of the Deaf,* 74, 389-395.

Jones, J. W. (17). "One Hundred Years of History in the Education of the Deaf in America and its Present Status." Paper prepared for Convention of American Instructors of the Deaf, Hartford, Conn.

Long, J. A. (31). *The Motor Ability of the Deaf.* To be published shortly.

Love, J. K. (96). *Deaf Mutism.* MacLehose and Sons, Glasgow.

Love, J. K. (12). "The Influence of Defects of Hearing in Relation to the Mental and Physical Development of the Child." *Child Study,* 5, 42-47.

MacMillan, D. P. and Bruner, F. G. (06). *Experimental Studies of Deaf Children.* Special Report of the Department of Child Study and Pedagogical Investigation, Chicago Public Schools.

McManaway, H. M. (23). "A Report of the Use of Standard Tests in the Virginia School for the Deaf." *Volta Rev.*, 25, 407-416.

Newlee, Clara E. (19). "A Report of Learning Tests with Deaf Children." *Volta Rev.*, 21, 216-223.

Pintner, R. (18). "A Measurement of Language Ability and Language Progress of Deaf Children." *Volta Rev.*, 20, 755-764.

Pintner, R. (18). "The Value of Mental Tests in the Classification of Pupils." *Am. Annals of the Deaf*, 63, 196-204.

Pintner, R. (28). "A Mental Survey of the Deaf." *J. of Ed. Psych.*, 19, 145-151.

Pintner, R. (29). "Speech and Speech-reading Tests for the Deaf." *J. of Appl. Psych.*, 12, 220-225.

Pintner, R. and Paterson, D. G. (15). "The Binet Scale and the Deaf Child." *J. of Ed. Psych.*, 6, 201-210.

Pintner, R. and Paterson, D. G. (15). "A Class Test with Deaf Children." *J. of Ed. Psych.*, 6, 591-600.

Pintner, R. and Paterson, D. G. (16). "Learning Tests with Deaf Children." *Psych. Rev.*, Monog., Vol. XX, No. 4. Whole No. 88.

Pintner, R. and Paterson, D. G. (16). "A Measurement of the Language Ability of Deaf Children." *Psych. Rev.*, 23, 413-436.

Pintner, R. and Paterson, D. G. (16). "The Survey of a Day-School for the Deaf." *Am. Annals of the Deaf*, 61, 417-433.

Pintner, R. and Paterson, D. G. (16). "The Ability of Deaf and Hearing Children to Follow Printed Directions." *Ped. Sem.*, 23, 477-497.

Pintner, R. and Paterson, D. G. (17). "Psychological Tests of Deaf Children." *Volta Rev.*, 19, 661-667.

Pintner, R. and Paterson, D. G. (17). "A Comparison of Deaf and Hearing Children in Visual Memory for Digits." *J. of Exper. Psych.*, 2, 76-88.

Pintner, R. and Paterson, D. G. (18). "Some Conclusions from Psychological Tests of the Deaf." *Volta Rev.*, 20, 10-14.

Reamer, J. C. (21). *Mental and Educational Measurements of the Deaf.* Psychological Monographs, No. 132, Princeton, N. J.

Upshall, C. C. (29). *Day Schools vs. Institutions for the Deaf.* T. C. Contribs. to Educ., 389.

CHAPTER XIX

THE BLIND

The First Intelligence Scale.—The blind, like the deaf, form a very small percentage of the total population, but their handicap is such as to require a very special method of education. The educators of the blind have, therefore, concentrated their attention upon methods of teaching and it is only within recent years that interest has been aroused in the general psychological make-up of the blind individual. One of the earliest workers in this field was himself a blind man, Mr. R. B. Irwin, Supervisor of the Education of the Blind in Cleveland, and his early interest in this work may best be described in his own words, as quoted by Hayes (20):—"In 1914 I went to Vineland with the idea of working with Dr. Goddard on a study of intelligence tests with a view to the revision of some existing scale so that it might be used more appropriately with the blind. It soon appeared that this was not a task that could be completed in one summer or one year. Dr. Goddard assisted me in the collection of all sorts of tests, which I arranged in a tentative order with a view to their being tried out on a large number of blind children. The aim was that they should ultimately be arranged and published in a form that would be useful to any mental examiner. The Binet measuring scale was taken as a basis. Tests which clearly depended upon vision for their proper performance were immediately stricken out. Other tests borrowed from anywhere and everywhere were added to the scale in the year in which they seemed most appropriate. The distribution of these tests was purely a matter of guesswork, and we realized it. After forty or fifty children in the Sunshine Homes were examined it was quite apparent that

some of our tests had been placed in the wrong year. I made very little change, however, as I felt that for a long time we must continue to collect data. The test sheets were printed with the view of testing the tests and not the subjects."

Continuing the work started by Irwin, Haines (16) modified and standardized his scale for the blind (see Chapter VI). Hayes (20) continued the work with the Irwin-Binet Scale for the blind. Hayes and Irwin (23) produced a revision of the Stanford-Binet, and Hayes (30) published a new revision of this scale for the blind.

Comparison of Blind and Sighted Pupils.—What little data are available seem to agree in finding the blind as a group inferior to the sighted on comparable intelligence tests. Hayes (20) makes a comparison between 670 blind children, tested by means of the Irwin-Binet tests, and 1,000 unselected children as tested by Terman.

	Percentage	
	Blind	*Sighted*
Genius	0.3	0.5
Very superior	1	2
Superior	5	9
Average	68	76
Dull	12	8
Borderline	7	2
Feebleminded	5	0.3

Pupils above average in intelligence are much less common among the blind than among the sighted. The greatest difference between the two distributions is found in the large percentages of dull, borderline, and feebleminded blind as compared with the sighted. The percentage of feebleminded blind is particularly large. Hayes (29) finds the median I.Q. of the blind below that of the sighted. When I.Q.s are calculated on those tests suitable for the blind, the median I.Q. is about 10 points below normal. This would put the median at about 90. Pintner's estimate for the deaf is about 80. These two estimates are at present the best that we have for the blind and

the deaf as contrasted with normal sighted hearing children.

In Hayes' results no differentiation is made between the partially blind and the totally blind. Haines (16) regards this distinction as important and arrives at the conclusion with respect to the totally blind that their intelligence "is not markedly inferior in grade, or different in quality from that of seeing subjects." On the other hand he finds the group of partially blind to be "more heavily laden with distinctly inferior mentalities." He believes "that some subnormality of vision has been made the excuse for constituting the school an asylum for some feebleminded persons who should be in institutions for the feebleminded. After eliminating such persons from the group, the average attainments of the remaining subjects, year by year, came very close to those of seeing subjects." Haines, however, recognizes that the amount of feeblemindedness in the ordinary institution for the blind, where no sharp distinction is made between totally and partially blind individuals, is much greater than the percentage of feeblemindedness in the general population. As in the case of the deaf, some causes of blindness are likewise causes of brain defect and, thus, we find many who are blind and feebleminded because of the same general inadequacy of the central nervous system. Hayes (23) has also studied the effect of the age of the child at which vision was lost. His age curve for the I.Q. is practically horizontal and hence he concludes that early loss of vision does not cause permanent mental retardation.

The evidence from group tests points also to the general inferiority of the blind. Hayes (20) reports that the average attainment of 122 blind subjects on the Pressey group tests is considerably lower than the average for the sighted. In a more extended and detailed study by Hayes (21) in which he used a great many intelligence and educational tests adapted for blind subjects, we note that, wherever adequate comparisons between blind and sighted subjects can be made, the blind are usually inferior at every age. This is true of the Trabue Completion Test, Terman Vocabulary Test, Pressey Practical Information,

Opposites, Analogies, Logical Judgment and Moral Judgment. Most of these comparisons are based on over 300 blind subjects in seven schools for the blind.

Practical Use of Tests.—Little has so far been reported as to the practical use of intelligence tests for classification purposes in schools for the blind, although undoubtedly something has been done that has not been published. Holterhoff (21) indicates that attention is being paid to the feebleminded blind and describes the type of education that should be undertaken with blind children whose I.Q.s are below 70. She suggests that the Revised Braille is better and easier to teach than the American Braille.

Sargent (24) makes a report of 500 former pupils of a blind school. The median I.Q.s on a scale for the blind for various groups lay between 89 and 98. About 41 per cent had I.Q.s below 90.

Special Mental Functions.—Much has been written upon the special ability of the blind with reference to memory, sensitivity of touch and hearing. The two latter functions do not concern us directly in this book, but it may be mentioned that Seashore and Ling (18) do not find the blind to be more sensitive or keen in sensory discrimination than seeing persons when fundamental capacities are tested, although they are undoubtedly superior in the general use of touch and hearing in practical life. With reference to memory, some have claimed superiority for the blind (Haines, 16) or equality of the blind and hearing (Bond and Dearborn, 17), but the more extended studies of Hayes (20) with unselected blind children seem to show that in general the blind are only slightly superior in rote memory, but show no superiority in logical memory. Special compensation for their visual defect either in touch, hearing or memory does not seem to exist.

Language and Educational Attainments.—Although in most educational tests the blind are somewhat inferior to the sighted (Hayes, 21), there does not seem to exist the great discrepancy in this respect, as we found to be the case among

the deaf. The language ability of the blind is much better than that of the deaf, and being better their whole educational development is easier and more rapid. Nevertheless, the reading of the blind is slow and difficult. Hayes (20) says, "The results as a whole indicate that under present conditions blind pupils attain at the end of their elementary school work a rate of reading only about one-third of that of sighted pupils of the fifth grade, and that they make no appreciable advance during the high school years." In other educational tests the blind are below the hearing. This is partly due to the presence of pupils of low mentality, according to Hayes (20), for he says, "Low mentality, then, affects the course of the curve somewhat, but when the low grade pupils are omitted, the curve is still far below the sighted standards."

Conclusions.—We may summarize this brief chapter by saying that a very good beginning has been made in the intelligence measurement of the blind. Enough research work has been done on the construction of scales, so that useful methods of measurement exist. The results of the tests so far published show the blind as a group somewhat inferior mentally to the sighted. There is evidently a large percentage of feeblemindedness among the blind, as we noted in the case of the deaf. Educational achievement among the blind is below the sighted, but probably superior to the deaf. The difficulty of acquiring language is not nearly as great among the blind as among the deaf, because the former learn to speak and talk easily and readily, just as a normal hearing-seeing child does. Lastly, there is no evidence that compensation for the handicap of blindness exists in the form of increased sensitivity of touch or hearing, or of better memory ability.

REFERENCES

Bond, N. J. and Dearborn, W. J. (17). "The Auditory Memory and Tactual Sensitivity of the Blind." *J. of Ed. Psych.*, 8, 21-26.

Haines, T. H. (16). *Mental Measurements of the Blind.* Psychological Monographs, Vol. XXI, No. 1. Whole No. 89.

Haines, T. H. (16). *A Mental Survey of the Ohio State School for the Blind.* The Ohio Board of Administration, Publication No. 9, 3-24.

Hayes, S. P. (20). "The Work of the Department of Psychological Research at the Pennsylvania Institution for the Blind, Overbrook." *Outlook for the Blind,* 14, 5-20.

Hayes, S. P. (20). "Mental and Educational Surveys in Seven Schools for the Blind." Report of Twenty-fifth Convention of American Association of Instructors of the Blind. Overlea, Maryland.

Hayes, S. P. (21). *Self-Surveys in Schools for the Blind.* Publications of the Pennsylvania Institution for the Instruction of the Blind, No. 2, pp. 1-60.

Hayes, S. P. (23). *Preliminary Study of the Influence upon School Success of the Age at Which Vision is Lost.* Publications of the Perkins Institution and Massachusetts School for the Blind, No. 1, Jan.

Hayes, S. P. (27). *Ten Years of Psychological Research in Schools for the Blind.* Publications of the Penna. Institution for the Blind, No. 4, Jan.

Hayes, S. P. (29). "The New Revision of the Binet Intelligence Tests for the Blind." *Teachers' Forum for Instructors of Blind Children,* 2, No. 2, 2-4.

Hayes, S. P. (30). *Terman's Condensed Guide for the Stanford Revision Adapted for the Blind.* Perkins Institution for the Blind, Watertown, Mass.

Hayes, S. P. and Irwin, R. B. (23). *Condensed Guide for the Stanford Revision Adapted for Use with the Blind.* Hayes, Mt. Holyoke.

Holterhoff, L. (21). "The Feebleminded Blind." *Sch. and Soc.,* 14, 174-179.

Sargent, R. F. (24). *What Can the Blind Do?* Publications of the Penna. Institution for the Instruction of the Blind. Overbrook, Philadelphia, No. 3, June.

Seashore, C. E. and Ling, T. L. (18). *The Comparative Sensitiveness of Blind and Seeing Persons.* Psychological Monographs, Vol. XXV, No. 2. Whole No. 108, 148-158.

CHAPTER XX

THE NEGRO

General Studies.—The universal interest in racial differences has led the psychologist to ask the question as to whether there exist differences in intelligence between different races as measured by our intelligence tests. The most extensive and most satisfactory comparison of any two races by means of intelligence tests has been made with reference to whites and negroes in this country. We shall, therefore, deal with this comparison in this chapter, and reserve for the next chapter the reports of tests of other racial groups.

There is an enormous amount of literature dealing with the negro in America and in many cases the difference between the negro and the white is pointed out. Writers range from a belief in practically no difference in intelligence between the negro and the white to a belief in a very profound difference. Almost always such opinions are based upon mere observation or hearsay knowledge or on studies of social conditions and the like. The work of Odum (10) is a good example of this type of study. He echoes the common opinion that negro children develop up to a certain age and then stop. He says "the brightest students are those from 9 to 13 years of age; the clearest minds seem to be found from 10 to 12 years of age." And again, "After 10 or 12 their development is physical rather than mental." It is, of course, very difficult to know precisely what is meant by such statements. Obviously a clear understanding of any differences that may exist can only be arrived at by careful and accurate measurements by means of standard tests.

The comparison of the school standing of negroes and whites

has been repeatedly made. The study of Mayo (13) is a good sample of this type of comparison. He studied high school students in New York City and found that 29 per cent of the colored reach or surpass the median school mark of the whites. He concluded that the colored are about three quarters as efficient as the white students. Furthermore, they are seven months older than the whites, and progress more slowly through the grades. The colored high school students are in all probability a more highly selected group than the white students. Only the brighter and more ambitious are likely to continue their studies in high school.

We shall confine ourselves in this chapter to a study of the results of intelligence testing and we shall disregard opinions and beliefs. We shall consider the test results under three headings, namely, Binet Test results, Group Tests, by which is meant the usual abstract verbal intelligence test, and other miscellaneous tests.

Binet Test Results.—Table 20 gives a summary of the most important results reported for the Binet Tests. We note in

TABLE 20

Comparison of Negro and White by Means of the Binet Tests

	Negro		*White*	
Author	*Median I.Q.*	*No. of Cases*	*Median I.Q.*	*No. of Cases*
Schwegler and Winn (20)	89	58	103	58
Arlitt (21)	83	71	106	191
Pintner and Keller (22)	88	71	95	249
Arlitt (22)	86	243	no white group	
Lacy (25-26)	91	817	103	5,159
Graham (26)	99	105	no white group	
Strachan (26)	93	609	102	14,463
Strachan (26)	92	375	101	6,063

every case that the negro I.Q. is below the white. The negro I.Q. ranges from 83 to 99 with a central tendency around 90.

Among other results that cannot readily be put into our table, we have the early work of Strong (13) using the Goddard Binet and making a comparison of 125 colored with 225 white children ranging in age from 5 to 15. The results are given in percentage retarded mentally:

	Colored	*White*
More than one year backward ..	29.4	10.2
Satisfactory	69.8	84.4
More than one year advanced ..	0.1	5.3

Again the author shows that the colored are more retarded even when compared with the poor whites as represented by the children of mill workers.

Lacy (18) in a study of 100 repeaters tested by the Binet Tests makes a comparison of the negroes and whites. At each age he finds the intelligence quotient of the whites superior to that of the negroes. He does not give the median I.Q.'s for the two groups, but from his distribution curve we can see that the median I.Q. of the negroes lies between 80 and 90, while that of the whites is probably between 90 and 100. This result for the colored group agrees with those previously quoted.

Negro college students would seem to test higher than negro elementary school children, just as we would expect, but in comparison with white students, they again fall lower. Derrick (20) gives 103 as the median I.Q. on the Binet for 55 colored students, and 112 as the median for 75 white students. The average chronological age of the white students is five years less than that of the negroes.

Group Tests.—In addition to these studies by means of the Binet, we have many reports of group tests. These group tests are for the most part the usual verbal abstract tests. These studies are numerous. We shall first consider the results of testing the negro draft in the army and then take up the tests of school children.

The best comparison by means of group tests of adults and, indeed, the widest and most convincing comparison of negroes

and whites in general is to be found in the army data. ("Memoirs," National Academy of Sciences, 21.) The mean mental ages and mean scores on the combined scale of intelligence for various samplings are as follows:

	Mean Score	*Mean M.A.*
White draft	13.5	13.1
Colored draft	9.9	10.4
White officers	18.8	17.3

The percentage of the colored reaching or exceeding the mean of the white draft is at most 12 per cent, and probably about 10 per cent. The combined scale is an evaluation of all scales used in the army.

The negro troops, however, score less than the white both on the Alpha and on the Beta scales. Alpha is a test for literates and Beta for illiterates, so that whether educated or uneducated, the negro is inferior to the white of similar status. The median scores are:

	Alpha	*Beta*
White draft-native	59	43
White draft-foreign	47	41
Colored draft-northern	39	33
Colored draft-southern	12	20

The percentage of the northern colored reaching or exceeding the median of the native white is on Alpha about 31, on Beta 34; while the percentage of the southern colored is on Alpha about 6, on Beta 14.

We may also study the difference between the racial groups by comparing the distributions of intelligence ratings, and by doing this we are reminded of the fact that there are negroes who receive the highest intelligence ratings and whites who receive the lowest, although the mode for the negroes is always further towards the lower end. The data below show the superiority of the white draft to the negro draft, the superiority of the white officer to the negro officer, and the superiority of the northern negro to the southern negro. For the significance

of the intelligence ratings, A, B, C, etc., the reader is referred to the tables in Chapter XIII. The figures show the percentages receiving the various intelligence ratings.

Drafted Men	*n*	*D—*	*D*	*C—*	*C*	*C+*	*B*	*A*
White	93,973	7.0	17.1	23.8	25.0	15.0	8.0	4.1
Negro	18,891	49.0	29.7	12.9	5.7	2.0	0.6	0.1
Officers								
White	1,385	0.1	0.3	0.7	6.2	12.3	31.2	49.2
Negro	95	3.3	10.0	5.3	22.1	21.0	24.2	14.7
Negro (training)	273	2.2	10.5	20.8	35.6	19.4	8.1	3.4
Section of Country								
Northern Negroes	4,705	14.4	31.2	25.8	18.0	7.2	2.7	0.7
Southern Negroes	6,848	57.0	29.2	9.6	3.4	0.7	0.2	0.1

From whatever angle, therefore, we look at these results, the marked superiority of the white group is apparent. The education of the negro has been for a long time and still is decidedly poorer than the education of the white, but the difference in intelligence is marked whether we compare the groups on tests for literates or illiterates. The discrepancy is also present when the results of the Binet tests of the two groups are compared. Ferguson (21) discusses the influence of education on the army tests with reference to negroes, and after making what he considers a reasonable allowance for difference in educational opportunity, he sums up by saying "that probably the safest and most reasonable expression of the relative intelligence of whites and negroes is that approximately 25 per cent of the latter equal or exceed the average of the former." This means, as Ferguson points out, a considerable amount of overlapping between the two groups. We are, however, warranted in saying that a real racial difference exists. The difference between the average white and the average negro in intelligence may not be very great, but it is quite definitely present, and this seems to be the case wherever two comparable groups of the two races have been examined, and even

TABLE 21

Negro and White School Children Tested by Group Tests

Author	n: Negro	n: White	Test	Negro I.Q.	Result
Pressey and Teter (19)	187	1,022	Pressey	. . .	Negro overlap 14%
Murdock (20)	225	514	Pressey	. . .	White + .85 P.E.
Haggerty (21)	3,000	13,000	Haggerty Delta	. . .	Negro overlap 14 to 20%
Jordan (22)	247	1,504	N. I. T.	75	Negro overlap 20 to 25%
Peterson (23)	37–734	71–641	Otis	58–75	Negro — 1.4 P.E.
			Pressey		
			Haggerty		
			Myers		
Sunne (24)	1,112	5,834	N. I. T.	. . .	Negro overlap from 14 to 33%
			Myers		
Garth and Whatley (25)	1,272	. . . . . .	N. I. T.	75	
St. Louis Report (25)	1,574	8,998	Pintner-Cunningham	92	White I.Q. 100
Hirsch (26)	449	1,030	Pintner-Cunningham and Dearborn	85	American White I.Q. 98
Davis (28)	222	. . . . . .	Terman	78	
Peterson and Lanier (29)	86	119	3 group tests	. . .	Negro — 2.2Q
	86	119	3 ingenuity tests	. . .	Negro — 0.8Q
Kempf and Collins (29)	399	677	Various	71	White I.Q. 88
Young (29)	314	323	N. I. T.	. . .	Whites 95% better
Gray and Bingham (29)	258	219	Not given	76	White I.Q. 108
Garth et al. (30)	2,006	. . . . . .	Otis	76 or 78	Negro overlap by age from 1 to 32

when we make allowance for the differences in educational opportunity.

Let us now examine the results of various group tests of negro and white children. Table 21 gives a summary of the results. In nine reports an I.Q. for the negro has been calculated. These range from 58 to 92 with a central tendency around 76, much lower than the central tendency for the Binet results. Five reports express their results in terms of the percentage of negroes reaching or exceeding the median of the whites. For purposes of brevity we have called this an overlap in our table. These percentages range from 1 to 33. The third method of comparing the two races is to express the difference between the means or medians in terms of the probable error or the semi-interquartile range (Q). We note considerable differences between the two racial groups when this method is employed. All of this evidence, therefore, supports the results found in the army. Negro children are much inferior to white children on our group intelligence tests.

The Peterson and Lanier (29) study shows the great difference between negro groups in various sections of the country. Nashville negroes are much inferior to Chicago and New York negroes. The latter are superior to Nashville whites on the rational learning test, and there is practically no difference between New York negroes and New York whites living in the same neighborhood. The authors comment on the severe selection of New York negroes, which may result in a negro group above the white in intelligence because of the severe struggle for existence.

There still remain a few reports on high school and college students. Thorndike (23) compares the two groups on the I.E.R. tests. The average scores by grade are:

	Negro		*White*	
	Score	*n*	*Score*	*n*
Grade IX	227	151	393	1,656
Grade X	280	114	413	492
Grade XI	296	84	426	505

Less than 4 per cent of the colored reach or exceed the median white score of the corresponding grade. Here we have a difference between the two races much greater than any so far reported. This is all the more significant because the author shows that the selection of the negro children in high school is a little narrower and so presumably a little better qualitatively than that of the whites. The comparison here is between children in a large city of the North Central division. If we attempt to explain why this difference is greater than that in the army or in the elementary school, we can only suggest that the tests used are perhaps more abstract than the others that have been used. But we need more results in order to test this hypothesis, that the racial difference increases the more abstract the tests become.

There are two reports for college students. Price (29) gives the results for negro students in seven negro colleges and finds the percentage of overlap on the Otis test to range from 7 to 43. For the total group the white median score is 49, the negro 39, and the overlap 20. In terms of I.Q. the white is 109 and the negro 98. Here again we find the overlap very similar to that for school children. Graham (30) reports the results of many intelligence tests given to the students at Fisk University. In general the overlap is about 36 or 37. This is higher than most reports. Peterson and Lanier (29) compare a white and negro normal school and find differences in favor of the white students as follows:

Verbal group intelligence	1.56 Q.
Non-verbal group intelligence	1.00 Q.
Mechanical ability	0.98 Q.

Other Tests.—If we raise the question as to negro and white results on tests other than the usual Binet and verbal group tests, we find several reports.

Negro children were tested on several different tests by Ferguson (16) and only on one test, the cancellation test, did they exceed the whites. He concludes that the average performance

of the colored is only three quarters as efficient as that of the whites. He also believes that there is a decided correlation among the colored between intelligence and the amount of white blood.

Pyle (15) compared colored and white children on thirteen different group tests involving memory, learning ability, association, and so forth. He compares the sexes and different social groups. In general, he finds that "the marks indicating the mental ability of the negro are about two-thirds those of the white."

Goodenough's (26) results on her test of drawing a man give the following results:

	Median I.Q.	*n*
American Whites	100.3	500
Californian Negroes	82.7	69
Southern Negroes	76.5	613

These negro I.Q.s are comparable to those found on the usual group test with school children. Klineberg (28) compared 129 negroes with 25 whites on twelve of the Pintner-Paterson Performance Tests and found a median I.Q. of 83 for the negroes and 88 for the whites. He finds the superiority of the whites to be solely due to speed. Davenport and Steggerda (29) gave various performance and verbal tests to comparable groups of negroes and whites in Jamaica and found the whites superior to the negroes on all tests except the repetition of seven digits. McGraw (31) tested 68 white and 60 negro infants on the Bühler Tests and found the mean developmental quotient for the white infants to be 105 as compared with a mean D.Q. for the negroes of 92. Only 28 per cent of the negroes were above the median of the whites.

Differences in Specific Traits.—Many workers, including those we have cited above, have made attempts to find a qualitative difference in intelligence between the two races. With some the motive seems to be to show that the quantitative difference in intelligence between negroes and whites is due to the

fact that the intelligence scales used have favored the type of thing which gives a preference to the specific traits in which the white race are superior. In other workers, this motive is not present and they have merely pointed out the traits in which the white and colored are equal, or in which the colored are superior to the white, or they have made a comparison of the colored themselves with regard to different traits. The conclusions to be drawn from the evidence presented are conflicting and difficult to interpret. Let us list some of the suggestions made by different workers:

The negro is better in concrete problems than in abstract. (Derrick, 20.)
The two races are equal in rote memory. (Schwegler and Winn, 20.)
In tests involving common sense adjustment to practical situations of a familiar type the colored are equal to the white. (Schwegler and Winn, 20.)
In tests of abstract reasoning colored are inferior to white. (Schwegler and Winn, 20.)
The colored children do best in a test of rote memory. (Pressey and Teter, 19.)
The colored do poorest in tests involving a knowledge of abstract terms. (Pressey and Teter, 19.)
Negroes rate lower than whites or Indians on a scale of resistance to fatigue. (Garth, 20.)
Negroes are less accurate. (Baldwin, 13.)
Negro children not superior to white in verbatim reproduction or immediate retention. (Sunne, 17.)
Negroes show a greater facility in control of words, a more fertile imagination as relating to general human activities, and a more original and perhaps more primitive taste in use of colors. (Sunne, 17.)
Negroes are not so speedy on performance tests (Klineberg, 28).
Negroes are about 1 1-3 times as suggestible as whites (Young, 29).

With reference to musical tests we have the following three conflicting reports:

Whites superior. (Gray and Bingham, 29.)
Negroes superior. (Davenport and Steggerda, 29.)
69 per cent of whites above negroes. (Peterson and Lanier, 29.)

Lastly in the "Memoirs" of the Academy of Sciences (21), we have a report based on the Army Tests, as follows: "The report takes the above results to indicate that the negro as compared with the white man of equal intelligence is relatively strong in the use of language, in acquaintance with verbal meanings, and in perception and observation; and that he is relatively weak in judgment, in ability to analyze and define exactly, and in reasoning." Some workers in the army and some other investigators fail to find any clearly marked qualitative differences.

A survey of this evidence, in parts conflicting, and as a whole couched in rather vague and general terms, leads one inevitably to the conclusion that we have not yet found with regard to the tests used any clearly marked qualitative differences between the white and the negro. Chance samplings of whites compared with each other might show the same slight differences which have been recorded as existing between negroes and whites. While we cannot conclude dogmatically that there are no qualitative differences between the two races, we may at least say that any marked qualitative difference in intelligence, as measured by the tests under consideration, does not seem to exist. In saying this we must remember that we are restricting ourselves almost entirely to the type of intelligence which is largely verbal or abstract, that is the sort of thing which is tested most effectively by the Binet and customary group tests. Mechanical intelligence has been much less effectively measured and social intelligence practically not at all. Furthermore we must remember that we have not raised the question as to emotional and moral differences between the two races. Popular opinion assumes large differences in these traits. There are no scientific comparisons of these as yet, and speculation about possible differences in this regard does not belong in a book devoted to intelligence testing.

Conclusions.—Our conclusions can be briefly summarized by saying that all results show the negro decidedly inferior to the white on standard intelligence tests. This difference is present among infants, elementary school children, high school pupils, university students and adult men. It occurs on Binet tests, on verbal group tests, on non-verbal group tests and on performance tests. These results are sufficiently numerous and consistent to point to a racial difference in intelligence. The overlapping of the two races is great, and the most liberal estimate seems to be that at most 25 per cent of the colored reach or exceed the median intelligence of the whites. No qualitative difference in intelligence between the two races can explain this marked quantitative difference. Indeed, it is doubtful whether real qualitative differences exist with reference to the traits measured by our intelligence tests.

REFERENCES

Arlitt, A. H. (21). "On the Need for Caution in Establishing Race Norms." *J. of Appl. Psych.,* 5, 179-183.

Arlitt, A. H. (22). "The Relation of Intelligence to Age in Negro Children." *J. of Appl. Psych.,* 6, 378-384.

Baldwin, B. T. (13). "The Learning of Delinquent Adolescent Girls as Shown by a Substitution Test." *J. cf Ed. Psych.,* 4, 317-332.

Davenport, C. B. and Steggerda, M. (29). *Race Crossing in Jamaica.* Carnegie Institution of Washington, Publication No. 395.

Davis, R. A. (28). "Some Relations between Amount of School Training and Intelligence among Negroes." *J. of Ed. Psych.,* 19, 127-130.

Derrick, S. M. (20). "A Comparative Study of the Intelligence of Seventy-five White and Fifty-five Colored College Students by the Stanford Revision of the Binet-Simon Scale." *J. of Appl. Psych.,* 4, 316-329.

Ferguson, S. O. (16). "The Psychology of the Negro." *Arch. of Psych.,* No. 36.

Ferguson, S. O. (21). "The Mental Status of the American Negro." *Sci. Monthly,* 12, 533-543.

Garth, T. R. (20). "Racial Differences in Mental Fatigue." *J. of Appl. Psych.,* 4, 235-244.

Garth, T. R. et al (30). "The Intelligence and Achievement of Southern Negro Children." *Sch. and Soc.,* 32, 431-435.

Garth, T. R. and Whatley, C. A. (25). "The Intelligence of Southern Negro Children." *Sch. and Soc.,* 22, 501-504.

Goodenough, F. L. (26). "Racial Difference in the Intelligence of School Children." *J. of Exp. Psych.*, 9, 388-397.

Graham, V. T. (26). *Health Studies of Negro Children.* I. Intelligence Studies of Negro Children in Atlanta, Georgia. Pub. Health Rep., 41, 2759-2783.

Graham, J. L. (30). "A Quantitative Comparison of Rational Responses of Negro and White College Students." *J. of Soc. Psych.*, 1, 97-121.

Gray, C. T. and Bingham, C. W. (29). "A Comparison of Certain Phases of Musical Ability of Colored and White Public School Pupils." *J. Ed. Psych.*, 20, 501-506.

Haggerty, M. E. (21). *Virginia Public Schools.* Part II, Educational Tests. New York.

Hirsch, N. D. M. (26). "A Study of Natio-Racial Mental Differences." *Genetic Psych. Monog.*, 1, Nos. 3 and 4, Clark Univ.

Jordan, A. H. (22). "Notes on Race Differences." *Sch. and Soc.*, 16, 503-504.

Kempf, G. A. and Collins, S. D. (29). "A Study of the Relation Between Mental and Physical Status of Children in Two Counties in Illinois." *U. S. Public Health Reports*, 44, 1743-1784.

Klineberg, O. (28). "An Experimental Study of Speed and Other Factors in Racial Differences." *Arch. of Psych.*, No. 93.

Lacy, L. D. (25-26). "Relative Intelligence of White and Colored Children." *Elem. Sch. J.*, 25, 542-546.

Lacy, W. I. (18). "A Study of One Hundred Retarded Fourth Grade Pupils Tested by the Binet Scale." *Psych. Clin.*, 12, 16-24.

Mayo, M. J. (13). "The Mental Capacity of the American Negro." *Arch. of Psych.*, No. 28.

McGraw, M. (31). *A Comparative Study of a Group of Southern White and Negro Infants.* To be published shortly.

Memoirs (21). The National Academy of Sciences, Vol. XV, Part III, Chaps. 8 and 10.

Murdock, M. (20). "Study of Race Differences in New York City." *Sch. and Soc.*, 11, 147-150.

Odum, H. W. (10). *Social and Mental Traits of the Negro.* Studies in History, Economics and Public Law, Whole No. 99. Columbia University.

Peterson, J. (23). "The Comparative Abilities of White and Negro Children." *Comp. Psych. Monog.*, Vol. 1, No. 5.

Peterson, J. and Lanier, L. H. (29). "Studies in the Comparative Abilities of Whites and Negroes." *Mental Measurement Monog.* Serial No. 5.

Pintner, R. and Keller, R. (22). "Intelligence Tests of Foreign Children." *J. of Ed. Psych.*, 13, 214-222.

Pressey, S. L. and Teter, G. F. (19). "A Comparison of Colored and White Children by Means of a Group Scale of Intelligence." *J. of Appl. Psych.*, 3, 277-282.

Price, J. S. C. (29). "The Intelligence of Negro College Freshmen." *Sch. and Soc.*, 30, 749-754.

Pyle, W. H. (15). "The Mind of the Negro Child." *Sch. and Soc.*, 1, 357-360.

Report of St. Louis Public Schools (25), *Public School Messenger*, 23, 45-57.

Schwegler, R. A. and Winn, E. (20). "A Comparative Study of the Intelligence of White and Colored Children." *J. of Ed. Res.*, 2, 838-848.

Strachan, L. (26). "Distribution of Intelligence Quotients of Twenty-two Thousand Primary School Children." *J. of Ed. Res.*, 14, 169-177.

Strong, A. C. (13). "Three Hundred Fifty White and Colored Children Measured by the Binet-Simon Measuring Scale of Intelligence." *Ped. Sem.*, 20, 485-515.

Sunne, D. (17). "A Comparative Study of White and Negro Children." *J. of Appl. Psych.*, 1, 71-83.

Sunne, D. (24). "Comparison of White and Negro Children in Verbal and Non-Verbal Tests." *Sch. and Soc.*, 19, 469-472.

Thorndike, E. L. (23). "Intelligence Scores of Colored Pupils in High Schools." *Sch. and Soc.*, 18, 569-570.

Young, P. C. (29). "Intelligence and Suggestibility in Whites and Negroes." *J. Comp. Psych.*, 9, 339-359.

CHAPTER XXI

THE FOREIGN-BORN

Conflicting Opinions.—In none of our topics are the data so unsatisfactory and opinions so conflicting as for this topic of racial differences. Apart from the comparison of the negro and white races in America, which we have discussed in Chapter XX, we do not have any real comparisons of the intelligence of different racial groups. And we have entitled this chapter "The Foreign-Born," rather than "Racial Differences in Intelligence," to emphasize the fact that the chief data we possess refer to racial groups in this country.

That there are differences in intelligence and in other characteristics between different races, has been assumed by many anthropologists and psychologists writing on this topic. But the opinions as to the amount of such differences vary from practically zero to a fairly large amount. Anthropologists, like Deniker and Tylor (81), assume "lower" and "higher" races, and seem to take for granted differences in intelligence. Deniker assumes a correlation between the brain and intelligence. Tylor says that, "in measuring the minds of the lower races, a good test is how far their children are able to take a civilized education." Boas (11) is a good example of the anthropologist who minimizes the differences between races. The difference in intelligence between civilized and primitive races he believes to be very slight. Man's "faculties" are in general very much the same. Nevertheless, he concludes that negro intelligence is a little below the white. Le Bon (98) is a good example of the descriptive psychologist who, without any actual measurements, finds appreciable differences between races. In his attempt "to describe the psychological characteristics which con-

stitute the soul of races" he makes a scale of 4 steps: (1) the primitive races, such as the Fuegians, the aboriginal Australians; (2) the inferior, such as the Negro; (3) the average, such as the Chinese, Japs, Mongolian and Semitic races; (4) the superior, such as the Indo-European races. The difference between the groups refers both to character and intelligence. The groups differ, he maintains, in capacity to reason, in power of attention and observation, in ability to foresee consequences, and the like.

Let the above serve as samples of psychological and anthropological opinion as to racial differences uninfluenced by the results of actual measurement. Without adequate psychological tests of racial groups, there is bound to be a difference of opinion and further argument seems rather futile.

Sensory Tests.—The first comparative tests of different races were largely tests of sensory qualities. Rivers (04) compared the Todas in India with Englishmen in respect to visual acuity, visual illusions, tactile discrimination, sensitivity to pain, and the like, and found no real differences between the two groups. "Pure sense acuity is much the same in all races." Myers (11) refers to the lack of racial differences found in vision and other sensory qualities by the Cambridge Anthropological Expedition, and jumps rather abruptly to the conclusion that the mind of the average European peasant is not essentially different from the "primitive" or "backward" races. Some difference in acuity of hearing between more primitive races and whites is found by Bruner (08) in his comparison by means of auditory tests between whites and Ainus, Patagonian Indians, and Pigmies. Woodworth (10) gave many sensory tests to various racial groups represented at the St. Louis Fair in 1904, and concluded that "on the whole the keenness of the senses seems to be about on a par in the various races of mankind." The importance for us of Woodworth's work lies, however, in the fact that he, for the first time, used a test that approximates more to a test of intelligence than any of the sensory and psychomotor tests that had, up to that time, been

used. He used the Sequin Form Board and found very little difference between the whites, Indians, Eskimos, Ainus, Filipinos and Singhalese; but he found that the Igorots, Negritos and Congo Pigmies were much poorer on this test as compared with the previously mentioned list of races. We should conclude, therefore, from the above studies that in all probability there are few, if any, marked racial differences in sensory qualities, with the suggestion of possible racial differences in intelligence indicated in the last experiment mentioned.

Intelligence Tests.—When we come to intelligence tests of racial groups, apart from the negro, we find a certain number of data in this country. Whatever conclusions we may arrive at, it is well to emphasize at the beginning that they can refer only to the samplings of the races found in the United States, and that we cannot infer anything at present as to the intelligence caliber of the race or nation to which these individuals belong. Furthermore, it will be apparent as we describe the results that it is extremely difficult to be sure that the groups tested are even a fair sampling of the racial group as found in the United States. Any particular city or community may for economic reasons attract a certain type of immigrant and the representatives of any particular racial group in that community may not be representative of the race as found in this country. Only by numerous investigations or by a very widespread survey can we be sure of getting a fair sampling of any racial group.

Another difficulty in making comparisons of racial groups and one that has retarded the work a great deal is the language factor. Until recently there were no adequate non-language tests available. If language tests are used the foreign-language-speaking group may be severely handicapped, even although the children are born in this country and speak English. Such at least is the contention of several workers. Walters (24) estimates a handicap of 6 to 8 months on the Stanford Binet for 13-year-old children from foreign-language-speaking homes. Mead (27) finds that the scores of Italian children on

the Otis test increase according to the amount of English spoken in the home. Pintner and Keller (22) believe the Binet is not adequate for those from foreign homes. Pintner (23) shows that the percentage of foreign children in a given school reaching or exceeding the median of the American children is only 37 on a verbal intelligence test, whereas it is 50 on a non-language test, indicating the same amount of intelligence of these two groups living in the same neighborhood and attending the same school. Wang (26) studies foreign students in a university and finds that the Chinese are handicapped by language in taking the usual intelligence test.

There are other workers who do not believe that our usual tests are a handicap to the school child born in this country of foreign parentage. Some are inclined to maintain that the facility with which he adapts himself to our language is an index of intelligence. Still others (Young, 22) are inclined to argue because verbal tests correlate higher with school marks than do non-verbal tests, that this indicates they are better measures of intelligence than non-verbal tests and hence they deny any language handicap. Giardini (23-24) found no difference by giving the Detroit First Grade Test in Italian. Most of these contentions do not seem to bear upon the real point at issue. Undoubtedly facility in acquiring the English language is an index of intelligence and a verbal test might show the relative intelligence of foreign children among themselves. We cannot, however, compare their scores with English-speaking groups, because there the language environment is so entirely different. Again a high correlation with school work merely shows that the language handicap enters into school work as well as into the intelligence test. The verbal intelligence test may be all right for school classification purposes, but it would still be inadequate for measuring the real differences in intelligence.

In our presentation we shall treat separately those nationalities for which we have a fair amount of data. Then we shall group together all other nationalities. All this will refer to

TABLE 22

Tests of Italian Children in This Country

Author	Italians		Americans		Test	Other Intelligence Ratings
	I.Q.	*n*	*I.Q.*	*n*		
Arlitt (21)	85	87	106	191	Binet	
Pintner and Keller (22)	84	313	95	249	Binet	
Brown (22)	78	51	....	....	Binet	
Young (22)	88	248	107	402	Army Alpha	
	96	246	109	393	Army Beta	
Pintner (23)	...	102	....	121	N. I. T.	Italian 36% overlap
	...	102	....	121	Pintner N-L.	Italian 43% overlap
Colvin and Allen (23)	91	50	92	50	Binet	
	76	50	85	50	N. I. T.	
Bere (24)	85	100	....	....	Binet	
Madsen (24)	79	16	....	....	Binet	
Feingold (24)	97	206	103	892	Mod. Army Alpha	H. S. Freshmen
	100	29	107	286	Mod. Army Alpha	H. S. Juniors
	96	16	100	264	Mod. Army Alpha	H. S. Seniors
Pintner (24)	...	218	....	653	Pintner N-L.	$\frac{\text{Diff.}}{\text{P.E. Diff.}} = 4.9$
Graham (25–26)	85	42	99	59	Binet	Clinic cases
Kirkpatrick (26)	83	96	104	94	Illinois	Italian 11% overlap
	...	96	....	94	Army Beta	Italian 19% overlap
Hirsch (26)	86	350	98	1,030	P—C. and Dearborn	
Goodenough (26)	89	456	101.5	500	Goodenough	
Mead (27)	...	276	....	160	Otis	
Rigg (28)	91	140	105	8,130	N. I. T.	Italian 7% overlap
Kempf and Collins (29)	87	31	107	....	Binet	

children in this country. We shall then deal with adults in this country as tested in the army; and then with a few results from abroad. In going over the data for children in this country—our most important data—we must keep in mind the two problems we have discussed above, namely, the inadequacy of language tests for foreign-language-speaking groups and the difficulty of getting a fair sampling of any racial group in this country.

Italians.—There are many studies of Italian children in this country. Table 22 gives a condensed summary of the chief findings. If we take all the I.Q.s just as they stand, we find there are 17 of them, with a range from 76 to 100 and a mid-point around 87. The 14 I.Q.s for Americans range from 85 to 109, with a mid-point around 102. Some would hold that this difference between 87 and 102 represents roughly the difference in the mentality of the two races. The writer would contend that this difference is probably exaggerated because of the predominance of verbal tests used. Wherever we have non-language and language tests given to the same groups we find the Italians drawing closer to the Americans on the non-language tests. The results of Young (22), Pintner (23) and Kirkpatrick (26) show this. Nevertheless, even on these tests, the Italian groups are below the Americans. If the median I.Q. of American children in general is 100, then that of the Italians is undoubtedly below this. In all probability it is not as low as 87. A rough estimate might place it between 90 and 95.

American Indians.—Table 23 gives a summary of the results of intelligence tests of American Indians. The I.Q.s are for the most part low, the majority being in the seventies and eighties. The Jamieson and Sandiford (28) report shows how the Indian I.Q. rises markedly when non-language and performance tests are used. It seems reasonable to suppose that the usual I.Q. on language tests is lowered because of language difficulty. Klineberg (27) maintains that the Indians are slower but more accurate than the whites. They are penalized

on performance tests by counting speed. But the Jamieson and Sandiford report does not show this. Many of the reports in our table discuss the difference between mixed and full-bloods,

TABLE 23

Tests of American Indians

Author	*Test*	*n*	*Results*
Rowe (14).......	Binet	....	94 per cent below age
Hunter et al. (21)	Otis	715	Indian score 83; white score 123
Garth (22)......	N. I. T.	....	Mixed-Blood above Full-Blood
Garth (25)......	N. I. T.	1,050	Median I.Q. 69
Fitzgerald et al. (26)..........	N. I. T.; Otis	98	Median I.Q. 87.5
Goodenough (26).	Goodenough	79	Median I.Q. 86
Garth (27)......	N. I. T.	765	Average I.Q. 74–77 for different groups
Garth and Garrett (28)..........	N. I. T.	2,256	Average I.Q. 70–91 for different groups
Garth et al. (28).	Otis Classif.	1,000	Median I.Q. 70
Jamieson and Sandiford (28).	N. I. T.	717	Median I.Q. 80
	Pintner-Cunningham		Median I.Q. 78
	Pintner Non-Language		Median I.Q. 97
	P-P. Performance		Median I.Q. 92

and they usually find that intelligence increases with increase in white blood.

Jews.—Table 24 summarizes the results for Jewish children. Two reports (Davies and Hughes, and Hughes) are for a sur-

TABLE 24

TESTS OF JEWISH CHILDREN

Author	*Test*	*n*	*Jewish I.Q.*	*Non-Jewish I.Q.*	*Other Results*
Murdoch (20)	Pressey	489	....	...	53% overlap
Feingold (24)	Mod. Alpha	518	103	103	H. S. Freshmen
		208	103	107	H. S. Juniors
		148	96	100	H. S. Seniors
Bere (24)	Binet	100	98	...	
Seago and Koldin (25)	N. I. T.	250	....	...	Better than Italian in same grade
Graham (25–26)	Binet, etc.	47	105	99	
Hirsch (26)	P.C. and Dearborn	75	103	98	Polish Jews
		627	99.5	98	Russian Jews
Goodenough (26)	Goodenough	55	106	100	
Davies and Hughes (27–28)	Northumberland	1,894	....	...	Jews exceed non-Jews by 10.5 I.Q. points
Hughes (28)	Northumberland	1,894	....	...	Jews exceed non-Jews at each age, 8–13
Rigg (28)	N. I. T.	445	103	105	

vey in three London schools. All the others are in this country. Of the ten comparisons between Jews and non-Jews, we find that six favor the Jews, three the non-Jews and one makes the two groups equal. On the whole, therefore, there seems to be a tendency for the Jewish child to do slightly better than the non-Jewish in the United States and in England, and this in

TABLE 25

TESTS OF MEXICAN CHILDREN IN THE U. S. A.

Author	*Test*	*n*	*Results*
Young (22)	Army Alpha	51	I.Q. 87
	Army Beta	53	I.Q. 96
Garth (23).....	National	307	Below mixed-blood Indians
Sheldon (24) ...	Binet	100	I.Q. 89
Paschal and Sullivan (25)....	Six performance tests	415	All tests below American norms
Goodenough (26)	Goodenough	367	I.Q. 88.5
Garth (28).....	N. I. T.	1,004	I.Q. 78
Garretson (28)..	N. I. T. and P. C.	117	Ave. I.Q. for grades, 70–97
	Myers	117	Ave. I.Q. for grades, 68–91

spite of the fact that many of the Jewish children come from non-English-speaking homes.

Terman (25) found that his group of gifted children with I.Q.s above 130 had 10.5 per cent Jewish stock, which is 100 per cent excess of Jewish stock as contrasted with the general population of the cities from which the children came. In Garrett's (28-29) study of 296 Columbia College Freshmen we find that the Jewish students make the highest mean score on

the Thorndike Test, the Moss Social Intelligence Test, and in college marks. All of these results, therefore, lead to the general conclusion that the Jewish group in this country is well up to, if not above, the American norm in intelligence.

TABLE 26

TESTS OF CHINESE AND JAPANESE CHILDREN

Author	*Test*	*n*	*Results*
Chinese			
Yeung (21)....	Binet	109	I.Q. 97
Symonds (24) .	Pintner Non-Lang.	513	I.Q. 99
Murdoch (25).	N. I. T., Beta	23–61	50% overlap on non-verbal
Sandiford and Kerr (26)...	P.P. Performance	224	I.Q. 107
Graham (26)..	Binet	62	I.Q. 87
	Short Binet	73	I.Q. 91
Porteus and Babcock (26)	Mod. Binet	212	I.Q. 87
Japanese			
Fukuda (25) ..	Binet	43	I.Q. 98
Murdoch (25) .	N. I. T., Beta	23–61	50% overlap on non-verbal
Darsie (26)....	Binet	658	I.Q. 90
	Army Beta	...	No difference
Sandiford and Kerr (26)...	P.P. Performance	276	I.Q. 114
Porteus and Babcock (26)	Mod. Binet	229	I.Q. 85

Whether this superiority is true for Jews in Europe in general as compared with non-Jews, we do not know. If so, it is interesting to speculate as to how much of it is due to the selection of the fittest resulting from the severe persecution for many centuries.

Mexicans.—The results for various groups of Mexican children tested in this country are shown in Table 25. The general picture is one of low mentality. And this seems to be true for verbal, non-verbal, non-language and performance tests. Undoubtedly the recent Mexican immigration into this country is bringing in individuals of poor mental calibre.

Chinese and Japanese.—Tests of Chinese and Japanese children with American tests in the United States, Canada and Hawaii are shown in Table 26. If we study this table we notice that the average I.Q.s on the Binet range from 85 to 98, but on all non-language and performance tests the Chinese and Japanese are equal to or above the American norms. It would look as if a language handicap were present on the Binet Test. Our general conclusion would seem to be that the Chinese and Japanese samples that have been tested differ little, if at all, in intelligence from the American standard.

Other Racial Groups.—With reference to other national groups our data are so few as not to warrant separate tables for each nationality. The results have been summarized in Table 27. This table cannot claim to be complete, because results for national groups with very small numbers have been disregarded. Some studies of racial groups do not lend themselves to this type of tabulation and they have been omitted. But the table does represent the main results up to date.

As we look over the table we may make certain comments. The English and Scotch groups seem slightly above the average, the Irish slightly below. The German seem about at the average. The Swedish and Norwegian are again about average. The I.Q.s for this northern European group range from 95 to 109 and the central tendency would seem to be about 100 to 103. The two French Canadian I.Q.s are poor. The two Bohemian I.Q.s are about average. The I.Q.s for the Polish group are rather low. The two Finnish I.Q.s are fair. The Portuguese I.Q.s are all very low, except the one for the non-language test. About the other national groups it is useless to try to generalize. Almost all the tests used are verbal in type. Where

TABLE 27

TESTS OF MISCELLANEOUS RACIAL GROUPS

Author	*Test*	*n*	*Results*
English and Scotch			
Brown (22)	Binet	90	I.Q. 102
Feingold (24)	Modified Alpha	17–76	I.Q. 103–109 (H. S. groups)
Hirsch (26)	Pintner-Cunn. and Dearborn	213	I.Q. 101
Kempf and Collins (29)	Binet	106	I.Q. 105
Irish			
Feingold (24)	Modified Alpha	44–278	I.Q. 97–100 (H. S. groups)
Hirsch (26)	Pintner-Cunn. and Dearborn	214	I.Q. 96
German			
Brown (22)	Binet	67	I.Q. 102
Pintner (23)	N. I. T.	45	36% overlap
	Pintner Non-Language	45	62% overlap
Pintner (24)	Pintner Non-Language	161	No diff. from Amer. scores
Feingold (24)	Modified Alpha	13–86	I.Q. 99–105 (H. S. groups)
Fukuda (25)	Binet	25	I.Q. 96
Hirsch (26)	Pintner-Cunn. and Dearborn	190	I.Q. 98.5
Rigg (28)	N. I. T.	1095	I.Q. 105
Kempf and Collins (29)	Binet	713	I.Q. 98
Swedish			
Brown (22)	Binet	187	I.Q. 102
Feingold (24)	Modified Alpha	28–114	I.Q. 98–102 (H. S. groups)

Author	*Test*	*n*	*Results*
Swedish—Continued			
Fukuda (25).....	Binet	63	I.Q. 95
Hirsch (26)......	Pintner-Cunn. and Dearborn	236	I.Q. 102
Swedish and Norwegian			
Kempf and Collins (29)......	Binet	113	I.Q. 104
French			
Brown (22)......	Binet	199	I.Q. 95
French Canadian			
Kirkpatrick (26) .	Illinois	158	I.Q. 87
Hirsch (26)......	Pintner-Cunn. and Dearborn	243	I.Q. 85
Bohemian			
Bere (24)........	Binet	100	I.Q. 93
Rigg (28)........	N. I. T.	118	I.Q. 104
Slavish			
Pintner and Keller (22).......	Binet	130	I.Q. 85
Polish			
Pintner (23).....	N. I. T.	18	41% overlap
	Pintner Non-Language	18	61% overlap
Feingold (24)....	Modified Alpha	11–90	I.Q. 92–100 (H. S. groups)
Fukuda (25).....	Binet	27	I.Q. 83
Hirsch (26)......	Pintner-Cunn. and Dearborn	227	I.Q. 90
Kempf and Collins (29).......	Binet	33	I.Q. 89
Lithuanian			
Hirsch (26)......	Pintner-Cunn. and Dearborn	468	I.Q. 97
Finnish			
Brown (22)......	Binet	226	I.Q. 90
Kirkpatrick (26) .	Illinois	147	I.Q. 100

Author	*Test*	*n*	*Results*
Hungarian			
Pintner and Keller (22).......	Binet	99	I.Q. 89
Portuguese			
Young (22)......	Binet	23	I.Q. 84
	Army Alpha	77	I.Q. 87
	Army Beta	75	I.Q. 96
Porteus and Babcock (26)......	Modified Binet	105	I.Q. 84
Hirsch (26)......	Pintner-Cunn. and Dearborn	671	I.Q. 83
Greek			
Hirsch (26)......	Pintner-Cunn. and Dearborn	270	I.Q. 88
Armenian			
Goodenough (26)	Goodenough	123	I.Q. 92
Filipinos			
Carreon (26).....	Haggerty	665	Much below Amer. norms
Hawaiians			
Porteus and Babcock (26)......	Modified Binet	105	I.Q. 85

non-language tests have been used we frequently have the suggestion of better relative performance on such tests.

Army Results.—Some confirmation of the results above is given by the army comparison of racial groups. Here we are dealing with adults tested by means of group tests. In this case, the men are all foreign-born, whereas with the children most of the cases are American-born children of foreign parentage. The army results are expressed in terms of the combined scale, which is a composite of the Alpha and Beta tests. Men who could not read and write English were given the Beta, a non-verbal test. The results given below show the average score, the equivalent mental age and the number of cases.

	Canada	*British Isles*	*Denmark, Norway, Sweden*	*Germany, Austria*	*Greece*	*Russia*	*Italy*
Mean score	13.7	13.4	13.3	13.2	11.9	11.2	11.0
Mental Age	13.3	13.0	12.9	12.9	11.9	11.3	11.2
No. of cases	948	1,214	1,610	597	573	2,701	4,002

The difference between the northern European and the southern and eastern European group is marked.

The army data allow us to make a comparison of foreign groups with reference to the percentage of superior and inferior men. Below are given the countries ranked in order according to the percentage of men receiving intelligence ratings of A or B, in the second column the per cent of A or B men is given, and in the third column the per cent of inferior men, i.e., men receiving intelligence ratings of D, D — or E:

Country	*Per Cent A.B.*	*Per Cent D, D —, E*
England	19.7	8.7
Scotland	13.0	13.6
White draft	12.1	24.1
Holland	10.7	9.2
Canada	10.5	19.5
Germany	8.3	15.0
Denmark	5.4	13.4
Sweden	4.3	19.4
Norway	4.1	25.6
Ireland	4.1	39.4
All foreign countries	4.0	45.6
Turkey	3.4	42.0
Austria	3.4	37.5
Russia	2.7	60.4
Greece	2.1	43.6
Italy	0.8	63.4
Belgium	0.8	24.0
Poland	0.5	69.9

Here again the tendency for the men from northern Europe to be superior is apparent. Belgium is a notable exception. England and Scotland contribute a larger percentage of superior men than is found in the white draft in general. These two countries also contribute fewer inferior men than are found in the white draft. If the countries were ranked according to the percentage of inferior men, the rank order would be somewhat changed, but not materially. Belgium would rise considerably in the scale. The southern and eastern European countries contribute by far the largest percentages of inferior men. These percentages are very great for Russia, Italy and Poland.

The average scores on the combined scale for the army results are given by Brigham (23) as follows:

Country	*Average Score*	*n*
England	14.87	411
Scotland	14.34	146
Holland	14.32	140
Germany	13.88	308
Native-born whites	13.77	81,465
Denmark	13.69	325
Canada	13.66	972
Sweden	13.30	691
Norway	12.98	611
Belgium	12.79	129
Ireland	12.32	658
Austria	12.27	301
Turkey	12.02	423
Greece	11.90	572
Russia	11.34	2,340
Italy	11.01	4,009
Poland	10.74	382

So far as we can make any comparison between these results for adults born in foreign countries with the results for the children of foreign-born, we note a fair degree of agreement. Italian and Polish children rate low and so do the Italian and Polish men in the army. The English and Scotch rate high in

both comparisons. The Germans and Scandinavians rate about average in both cases. The army results unfortunately do not give any returns for Jews, Mexicans and American Indians.

Racial Comparisons Abroad.—All of the results discussed up to this point have been concerned with the racial samplings that have been found in this country. We must emphasize again the point that they do not give us any basis for the comparison of racial or nationality groups in general. To obtain adequate comparisons between two national groups, say between Americans in the United States and Italians, we need adequate samplings on the same test in both countries. The tests must be the same and be equally fair to the groups to be compared.

Tests involving language, when translated into various languages, can never form an adequate comparison. Non-language tests seem to the writer to offer the only possibility. So far, very little work along this line has been done. We may quote as a sample of this type of work the comparison of two Indian groups by Herrick (21) in India. He compared the performance of Panchama and Brahman children by means of the Goddard Form Board. The Panchamas belong to the lowest castes, the Brahman to the highest. The results show the Panchamas somewhat below the Brahmans at most ages, but the difference between the two groups is not large. Both groups seem inferior to the norms for American children.

A comparison of Chinese college students in China with American college students in this country has been made by Walcott (20). The Chinese were tested by means of English language tests as they were all thoroughly familiar with that language. On the Stanford Revision 44 out of 63 have I.Q.'s above 100. On a group test the median score for the Chinese is below that for the American students. It is obvious, however, that we cannot get any indication of racial differences in intelligence from such studies.

Pintner's (27) comparison of Belgian and American children comes nearest to the type of study necessary to compare na-

tional groups. The Pintner Non-Language Test was given in exactly the same way as given in this country to 271 Belgian school children. This small number of children naturally cannot be considered representative of Belgian children in general, but the school was described as an average Belgian school. The mean scores for ages 9 to 14 inclusive are practically the same as those for American children. No reliable differences were obtained.

Within Belgium itself, Walloon and Flemish children have been compared by Decroly (26) by means of a translation of the Ballard Test, an English group test. The number of children tested was 7,160, about half Flemish and half Walloon. The Walloons scored higher than the Flemish at all ages. Wood's (29) results of the Otis Higher S. A. Test at Constantinople Women's College shows practically no difference between small samples of Bulgarian, Armenian, Turkish and Greek students.

The Immigrant.—Although some mental testing of the immigrant has been done at various times, little of it has been published. Feeblemindedness is one reason for refusing admittance and a certain number are deported for this reason every year. No general intelligence test is given to all immigrants, and under existing conditions is very likely not practicable. The cases of feeblemindedness detected by the medical officers are in all probability of low grade. Detection is more likely to occur where mental inferiority exists along with some physical defect. The recent quota system has resulted in some testing by American officials abroad, before the immigrant sets out to this country.

The reports of Goddard (13, 17) are among the very few which we possess dealing with actual testing at an immigrant receiving station. Goddard believes that trained workers with the feebleminded could pick out by inspection as the immigrants walk past about 90 per cent of the feebleminded cases in contrast to only 10 per cent selected by the physicians. There is no evidence, however, for this belief.

Tests were made under Goddard's direction at the immigrant receiving station at Ellis Island, New York, in 1913. The tests used were the Binet Scale, as well as three or four performance tests. The testing was carried on by means of interpreters. The groups chosen for testing were selected in the sense that they were taken after the physicians had culled out all the obviously feebleminded. To offset this Goddard's workers passed over the "obviously normal." This left what Goddard calls the "great mass of average immigrants." This method of selecting cases makes it difficult to tell exactly what sort of groups were tested, and it makes it still more difficult to interpret the findings.

The following results are given in per cents:

	Normal	*Borderline*	*F.M.*	*No. of Cases*
Jews	10	7	83	30
Hungarians	0	20	80	20
Italians	7	15	79	48
Russians	0	9	87	43

All of the groups show an abnormally high percentage of feeblemindedness, and this is due in all probability to the method of selection and the standards used for interpreting the results of the tests. No one is willing to believe without much more adequate evidence that the percentages given above are representative of the intelligence caliber of the various immigrant groups.

However inconclusive these data are, they, nevertheless, raise the very practical question of the value of intelligence tests in the selection of the immigrant. We may say that the problem is now entirely one of administration or procedure, because there are now a sufficient number of individual performance tests and group non-language tests to enable examiners to give an adequate intelligence examination. The question as to what level of intelligence should bar the immigrant is one that could readily be established after some experimentation. Indeed,

the level might fluctuate somewhat according to the number of immigrants that the country might decide to admit each year, but a certain level might well be established below which none would be admitted because of their low mental capacity. Whether such tests should be given in the country of origin or at this side of the ocean are questions of administration involving many difficulties. It is well, however, to remember that we already have established the practice of refusing admission to immigrants because of mental defect. It remains, therefore, merely to conduct the selection by well-established psychological methods of intelligence measurement. Boody (26) discusses the practical problems to be met in such work.

We need here only indicate the social significance of this problem. Mental ability is inherited. The population of the United States is largely recruited by immigration. The country cannot afford to admit year after year large numbers of mentally inferior people, who will continue to multiply and lower the level of intelligence of the whole nation. Our tests, although inconclusive, would seem to indicate that the level of certain racial groups coming to this country is below that of the nation at large. Increased vigilance is, therefore, required. Literacy tests and restriction of groups based on quotas apportioned to the racial groups already here are all helpful and tend perhaps in the long run to raise the average intelligence level of the immigrant as compared with unrestricted immigration. But these checks are by no means sufficient. Indeed the last measure may at times be harmful, if it checks immigrants of high mental caliber coming for some reason or another from a country whose quota happens to be small, because for historic reasons that country does not happen to be well represented at present in the United States. It would be well, therefore, to emphasize the intelligence factor in the selection of our immigrants to a much greater extent than we do at present, lest the whole nation be diluted with stocks of inferior mentality. Intelligence is to be sure not the only factor to be considered in the selection of the immigrant. There are ques-

tions of health, physical well-being, moral qualities, questions of race, and the like. But intelligence is an important factor and one that has been sadly neglected up to the present time.

Conclusion.—No adequate racial or national comparisons by means of intelligence tests have yet been made. Such comparisons demand the use of the same test (not translations) with adequate samplings of the groups in their own countries. From this point of view, we know nothing about racial or national differences in intelligence. Our best samplings are United States whites with United States negroes where we have two national groups speaking the same language and, therefore, able to take the same tests. But we know nothing of the intelligence of the negro in general contrasted with the white. Many psychologists believe that there are racial differences in intelligence and shrewd inferences as to high and low races have been made, and may be correct, but, as yet, the scientific requirements for adequate racial comparisons have not been met.

The bulk of our results in this chapter deal with the children of parents of various national groups that have immigrated to this country. Most of this work is unfortunate because it is based upon the comparison of English-speaking American children with children speaking various amounts of a foreign language in their homes. Some psychologists believe this so-called language handicap to be negligible. Others believe that the results of language and non-language tests show it to be by no means negligible. Many of the results are further confused by the attempts of workers to allow for or counteract the language handicap by their own modifications of verbal tests. Private modifications of tests immediately spoil them as instruments of measurement. A modified test is an unstandardized test and we cannot interpret it.

Taking the results as they are, and bearing in mind their limitations, we may nevertheless venture some tentative conclusions. It would seem as if the forces attracting various national groups to this country have differed from group to group,

resulting in now a superior group from one country and now an inferior group from another. The northern European nations seem to be represented by a sampling in this country of relatively good intelligence; the southern European nations by relatively poor intelligence. Whether the sampling in this country represents the nations in Europe, we cannot tell. The national or racial samplings in this country that test high are the Jewish, the English and Scotch. Those that test around the average are the German, Scandinavian, Bohemian, Chinese and Japanese. Those that test low are the Italian, Polish, Portuguese, French Canadian, Mexican and American Indian. With reference to other groups our data are few.

The problem of testing the immigrant is an important question for a country like the United States, where thousands of future citizens are being admitted each year. There is enough poor intelligence in the country already which cannot be got rid of. It would, therefore, seem to be a wise policy to guard rigidly against the admittance of more of this low intelligence to a much greater extent than is apparently being done at present. We have a plethora of morons and dullards, but no excess of high intelligence.

REFERENCES

Arlitt, A. H. (21). "On the Need for Caution in Establishing Race Norms." *J. of Appl. Psych.*, 5, 179-183.

Bere, M. (24). *A Comparative Study of the Mental Capacity of Children of Foreign Parentage.* T. C. Contribs. to Educ., No. 154.

Boas, F. (11). *The Mind of Primitive Man.* New York.

Boody, B. M. (26). *A Psychological Study of Immigrant Children at Ellis Island.* Mental Measurement Monographs, No. 3, Baltimore.

Brigham, C. C. (23). *A Study of American Intelligence.* Princeton University Press.

Brown, G. L. (22). "Intelligence as Related to Nationality." *J. of Ed. Res.*, 5, 324-327.

Bruner, F. G. (08). "The Hearing of Primitive Peoples." *Arch. of Psych.*, No. 11.

Carreon, M. L. (26). *Philippine Studies in Mental Measurement.* Yonkers.

Colvin, S. S. and Allen, R. D. (23). "Mental Tests and Linguistic Ability." *J. of Ed. Psych.*, 14, 1-20.

Darsie, M. L. (26). *The Mental Capacity of American-Born Japanese Children.* Comparative Psychology Monog., Vol. 3, No. 15.

Davies, M. and Hughes, A. G. (27-28). "An Investigation into the Comparative Intelligence and Attainments of Jewish and Non-Jewish School Children." *Brit. J. of Psych.,* 18, 134-146.

Decroly, O. (26). "Essai d'application du test de Ballard dans les écoles belges." *Année Psych.,* 27, 57-93.

Deniker, J. (no date) *The Races of Man,* New York.

Feingold, G. A. (24). "Intelligence of the First Generation of Immigrant Groups." *J. of Ed. Psych.,* 15, 65-82.

Fitzgerald, J. A. and Ludeman, W. W. (26). "The Intelligence of Indian Children." *J. of Comp. Psych.,* 6, 319-328.

Fukuda, T. (25). "A Survey of the Intelligence and Environment of School Children." *Am. J. of Psych.,* 36, 124-139.

Garrett, H. E. (28-29). "Jews and Others." *Person, J.,* 7, 341-348.

Garretson, O. K. (28). "A Study of Causes of Retardation among Mexican Children in a Small Public School System in Arizona." *J. Ed. Psych.,* 19, 31-40.

Garth, T. R. (21). "The Results of Some Tests on Full and Mixed Blood Indians." *J. of Appl. Psych.,* 5, 359-372.

Garth, T. R. (22). "A Comparison of Mental Abilities of Mixed and Full Blood Indians on a Basis of Education." *Psych. Rev.,* 29, 221-236.

Garth, T. R. (23). "A Comparison of the Intelligence of Mexican and Mixed and Full Blood Indian Children." *Psych. Rev.,* 30, 388-401.

Garth, T. R. (25). "The Intelligence of Full Blood Indians." *J. of Appl. Psych.,* 9, 382-389.

Garth, T. R. (27). "The Intelligence of Mixed Blood Indians." *J. of Appl. Psych.,* 11, 268-275.

Garth, T. R. (28). "The Intelligence of Mexican School Children." *Sch. and Soc.,* 27, 791-794.

Garth, T. R. and Garrett, J. E. (28). "A Comparative Study of the Intelligence of Indians in United States Indian Schools and in the Public or Common Schools." *Sch. and Soc.,* 27, 178-184.

Garth, T. R. et al. (28). "A Study of the Intelligence and Achievement of Full Blood Indians." *J. of Appl. Psych.,* 12, 511-516.

Giardini, G. and Root, W. T. (23-24). "A Comparison of the Detroit First Grade Tests Given in Italian and English." *Psych. Clin.,* 15, 101-108.

Goddard, H. H. (13-14). "The Binet Tests in Relation to Immigration." *J. of Psycho-Asthenics,* 18, 105-107.

Goddard, H. H. (17). "Mental Tests and the Immigrant." *J. of Delinq.,* 2, 243-277.

Goodenough, F. L. (26). "Racial Differences in the Intelligence of School Children." *J. of Exp. Psych.,* 9, 388-397.

Graham, V. T. (25-26). "The Intelligence of Italian and Jewish Children in the Habit Clinics of the Massachusetts Division of Mental Hygiene." *J. of Ab. and Soc. Psych.,* 20, 371-376.

Graham, V. T. (26). "The Intelligence of Chinese Children in San Francisco." *J. of Comp. Psych.*, 6, 43-71.

Herrick, D. S. (21). "A Comparison of Brahman and Panchama Children in South India with Each Other and with American Children by Means of the Goddard Form Board." *J. of Appl. Psych.*, 5, 253-260.

Hirsch, N. D. (26). "A Study of Natio-Racial Mental Differences." *Genetic Psych. Monog.*, Vol. 1, Nos. 3 and 4.

Hsiao, H. H. (29). "The Mentality of the Chinese and Japanese." *J. of Appl. Psych.*, 13, 9-31.

Hughes, A. G. (28). "Jews and Gentiles." *Eug. Rev.*, 20, 89-94.

Hunter, W. S. and Sommermier, E. (21). "The Relation of Degree of Indian Blood to Score on the Otis Intelligence Test." *Psych. Bull.* 18, 91-92.

Jamieson, E. and Sandiford, P. (28). "The Mental Capacity of Southern Ontario Indians." *J. Ed. Psych.*, 19, 313-328.

Kempf, G. A. and Collins, S. D. (29). "A Study of the Relation Between Mental and Physical Status of Children in Two Counties in Illinois." *U. S. Public Health Reports*, 44, 1743-1784.

Kirkpatrick, C. (26). *Intelligence and Immigration.* Mental Measurement Monographs, No. 2. Baltimore.

Klineberg, O. (28). "An Experimental Study of Speed and Other Factors in Racial Differences." *Arch. of Psych.*, No. 93.

Le Bon, G. (98). *The Psychology of Peoples*, New York.

Madsen, I. N. (24). "Some Results with the Stanford Revision." *Sch. and Soc.*, 19, 559-562.

Mead, M. (27). "Group Intelligence Tests and Linguistic Disability among Italian Children." *Sch. and Soc.*, 25, 465-468.

Memoirs, (21). The National Academy of Sciences, Vol. XV, Part III, Chap. 6.

Murdock, M. (20). "Study of Race Differences in New York City." *Sch. and Soc.*, 11, 147-150.

Murdoch, K. (25). "A Study of Differences Found between Races in Intellect and Morality." *Sch. and Soc.*, 22, 628-632 and 659-664.

Myers, C. S. (11). *On the Permanence of Racial Mental Differences.* Papers on Inter-Racial Problems, Ed. Spiller. The World's Peace Foundation, 73-79.

Paschal, F. C. and Sullivan, L. R. (25). "Racial Differences in the Mental and Physical Development of Mexican Children." *Comp. Psych. Monog.*, Vol. 3, No. 14.

Pintner, R. (23). "Comparison of American and Foreign Children on Intelligence Tests." *J. of Ed. Psych.*, 14, 292-295.

Pintner, R. (24). "Results Obtained with the Non-Language Group Test." *J. of Ed. Psych.*, 15, 473-483.

Pintner, R. (27). "Non-Language Tests in Foreign Countries." *Sch. and Soc.*, 26, 374-376.

Pintner, R. and Keller, R. (22). "Intelligence Tests of Foreign Children." *J. of Ed. Psych.*, 13, 214-222.

Porteus, S. D. and Babcock, M. E. (26). *Temperament and Race.* Boston.
Rigg, M. (28). "Some Further Data on the Language Handicap." *J. Ed. Psych.,* 19, 252-256.
Rivers, W. H. R. (04-05). "Observations on the Senses of the Todas." *Brit. J. of Psych.,* 1, 321-396.
Rowe, E. C. (14). "Five Hundred Forty-seven White and Two Hundred Sixty-eight Indian Children Tested by the Binet-Simon Tests." *Ped. Sem.,* 21, 454-468.
Sandiford, P. and Kerr, R. (26). "Intelligence of Chinese and Japanese Children." *J. of Ed. Psych.,* 17, 361-367.
Seago, D. W. and Koldin, T. S. (25). "A Comparative Study of the Mental Capacity of Sixth Grade Jewish and Italian Children." *Sch. and Soc.,* 22, 564-568.
Sheldon, W. H. (24). "The Intelligence of Mexican Children." *Sch. and Soc.,* 19, 139-142.
Symonds, P. M. (24). "The Intelligence of Chinese in Hawaii." *Sch. and Soc.,* 19, 442.
Terman, L. M. et al. (25). *Genetic Studies of Genius.* Stanford University Press.
Tylor, E. B. (81). *Anthropology.*
Walcott, G. D. (20). "The Intelligence of Chinese Students." *Sch. and Soc.,* 11, 474-480.
Walters, F. C. (24). "Language Handicap and the Stanford Revision of the Binet-Simon Tests." *J. of Ed. Psych.,* 15, 276-284.
Wang, S. L. (26). "A Demonstration of the Language Difficulty Involved in Comparing Racial Groups by Means of Verbal Intelligence Tests." *J. of Appl. Psych.,* 10, 102-106.
Wood, M. M. (29). "Mental Test Findings with Armenian, Turkish, Greek and Bulgarian Subjects." *J. of Appl. Psych.,* 13, 266-273.
Woodworth, R. S. (10). "Racial Differences in Mental Traits." *Science, N. S.,* 31, 171-186.
Yeung, K. T. (21). "The Intelligence of Chinese Children in San Francisco and Vicinity." *J. of Appl. Psych.,* 5, 267-274.
Young, K. (22). *Mental Differences in Certain Immigrant Groups.* Univ. of Oregon Publication, Vol. I, No. 11, Eugene, Oregon.

CHAPTER XXII

THE EMPLOYEE

The Value of Intelligence Tests.—The applications of psychology to industry and to the problems of vocational advice and guidance are far-reaching and complicated. The use of intelligence tests in this connection is merely one item in the problem. It is not the purpose of this chapter to attempt a résumé of the whole field. We shall be concerned only with the use of intelligence tests with the employee or prospective employee.

In the field of business and industry the psychologist has contributed most in the direction of the selection of individuals for positions. Relatively little has been accomplished in the more difficult task of the selection of positions for individuals (Ayres, 13). If we consider merely the various kinds of tests constructed by the psychologist, disregarding rating scales, job analysis, personnel specifications and other useful devices and methods, we may roughly divide them into (1) trade tests; (2) specific ability tests; (3) general ability tests.

Trade tests determine the amount of knowledge of or skill in any particular job. They are essentially achievement tests, analogous to educational tests in the school. They are not prognostic. They measure the amount of accomplishment which the individual now possesses. These tests were developed in the army (Personnel System of the U. S. Army, 19) and they have been described at length by Chapman (21), and by Toops (21). Specific ability tests are such as attempt to measure those special abilities which are supposed to enter into any particular job. Sometimes these tests resemble very closely the actual operation, physical or mental, which the

worker goes through in the performance of the job. At other times the test is very different from the job, but the performance of both is supposed to involve the same kind of ability. The third type of test is the intelligence test proper, which is not a test of trade ability and makes no presuppositions as to resemblance with specific abilities required for specific jobs. It is with this last type of test that we shall deal more particularly in this chapter. Obviously there can be no sharp line drawn between the three types of tests and certainly not between the second and third. As we shall see, tests for the selection of various sorts of workers are very often a combination of the second and third types.

The value of a general intelligence test in the selection of workers will depend largely upon the aim or purpose of the testing. If the only purpose of the psychological test is to fill a particular job more or less satisfactorily, then we may agree in general with Viteles (21) when he says, "Such tests have a very limited place in industry. They are used in the first place to shut out from employment the feebleminded, those who are altogether unfit for any job in the plant by reason of deficiency in mentality. They are also used to select workers for jobs in which success depends, to a very great extent, upon a high level of general intelligence, namely, the executive positions. For selection for the great mass of skilled and semi-skilled jobs, for office jobs and simple clerical jobs, the general intelligence test cannot be used. The carpenter and the tool-maker, for example, must stand on approximately the same level with reference to general intelligence, but very different specific abilities are required for these trades." If, however, the employment manager conceives his function to be broader and more comprehensive than merely filling satisfactorily a specific job, we will agree heartily with Carney (19) when he says, "The modern employment manager is not satisfied when he has merely introduced the applicant to his first job in the organization. He feels that he must follow up this first placement by such transfers and promotions as are necessary to

bring out all the best that the man is capable of attaining. To do less than this is not only unfair to the individual, it is also poor business. For nothing makes for a loyal force of enthusiastic workers so much as a vigorous educational and promotional program. The fact that those in high places have risen from the ranks is the strongest argument that there is a future for men of ability in the organization, and men of ability cannot be held long at any wage where they feel they are working in a blind alley. Therefore, in a set of tests for shop clerks it is essential to introduce measures of general intelligence which will pick out the men in whom it is worth while to invest broad training with the hope that they can ultimately become understudies for executive positions. For while no correlation was discovered between general intelligence and success as routine time-clerks, considerable correlation appeared between general intelligence and success in positions of greater responsibility. Probably even a moron, if he possessed peculiar talent for that special work, might write time-tickets successfully. But the ability to organize work and to train subordinates requires intelligence which is readily recognizable through tests."

The value of intelligence tests will, therefore, vary according to what we are trying to achieve, whether it be the selection of workers for specific jobs, or the selection of high grade individuals to be trained for important positions, or, again, the elimination of low-grade material. The importance of this last point is stressed by Rossy (17), and he calls attention to the fact that low grade individuals will probably increase the number of industrial accidents, help to lower the output, perhaps produce an inferior product and certainly increase the percentage of labor turnover. The feebleminded and the psychopathic he regards as particularly dangerous. Mental examinations can eliminate these.

The Intelligence Level of Various Occupations.—The most comprehensive comparison of the intelligence level of men employed in various occupations is afforded by the results of the army intelligence tests ("Memoirs," 21). The distribu-

tion of intelligence ratings and the median scores for a great many different occupations have been computed. The following list shows occupations well represented in the army ranked according to the median scores. The intelligence rating of the median for each group is given. For the significance of these letter ratings the reader is referred to pages 318, 319. These data are based upon results for over eighteen thousand men.

Occupation	*Median Intelligence Rating*
Laborer	C —
Miner	C —
Teamster	C —
Barber	C —
Horseshoer	C
Bricklayer	C
Cook	C
Baker	C
Painter	C
Blacksmith	C
Carpenter	C
Butcher	C
Machinist	C
Hand riveter	C
Telegraph lineman	C
Pipefitter	C
Plumber	C
Toolmaker	C
Gunsmith	C
Mechanic	C
Auto-repairman	C
Auto-engine mechanic	C
Auto assembler	C
Ship carpenter	C
Telephone operator	C
Concrete construction foreman	C +
Stock-keeper	C +

Occupation	*Median Intelligence Rating*
Photographer	C +
Telegrapher	C +
Railroad clerk	C +
Filing clerk	C +
General clerk	C +
Army nurse	C +
Bookkeeper	C +
Dental officer	B
Mechanical draughtsman	B
Accountant	B
Civil engineer	B
Medical officer	B
Engineer officer	A

These results do not mean that all the men in any one group achieve the intelligence rating assigned to that group. In most occupations men of all grades of intelligence are to be found. Inspection of the detailed results of the army data reveals at once a great overlapping between groups. The spread of the middle fifty per cent of each group is considerable. This is seen in Figure 27, where the length of the bar shows the spread of the middle fifty per cent of the individuals in each group. The median intelligence of laborers, for example, is C —, or rather poor intelligence. The middle fifty per cent range from D through C — to C. In addition there was a lower group of twenty-five per cent who made ratings of D and D —, and an upper group also of twenty-five per cent who made ratings of C, C +, B and A. As a matter of fact, of the 1,453 laborers tabulated, 10.7 per cent made ratings of C +, 3.5 per cent ratings of B, and 0.62 per cent ratings of A. In a similar way we must interpret all the other occupations.

Nevertheless, it is obvious that a selective process is at work, which tends to prevent men of low intelligence from filling certain positions, and which favors the men of high intelligence

for certain positions. As Thorndike (19) puts it: "No less significant is the variability within each occupational group.

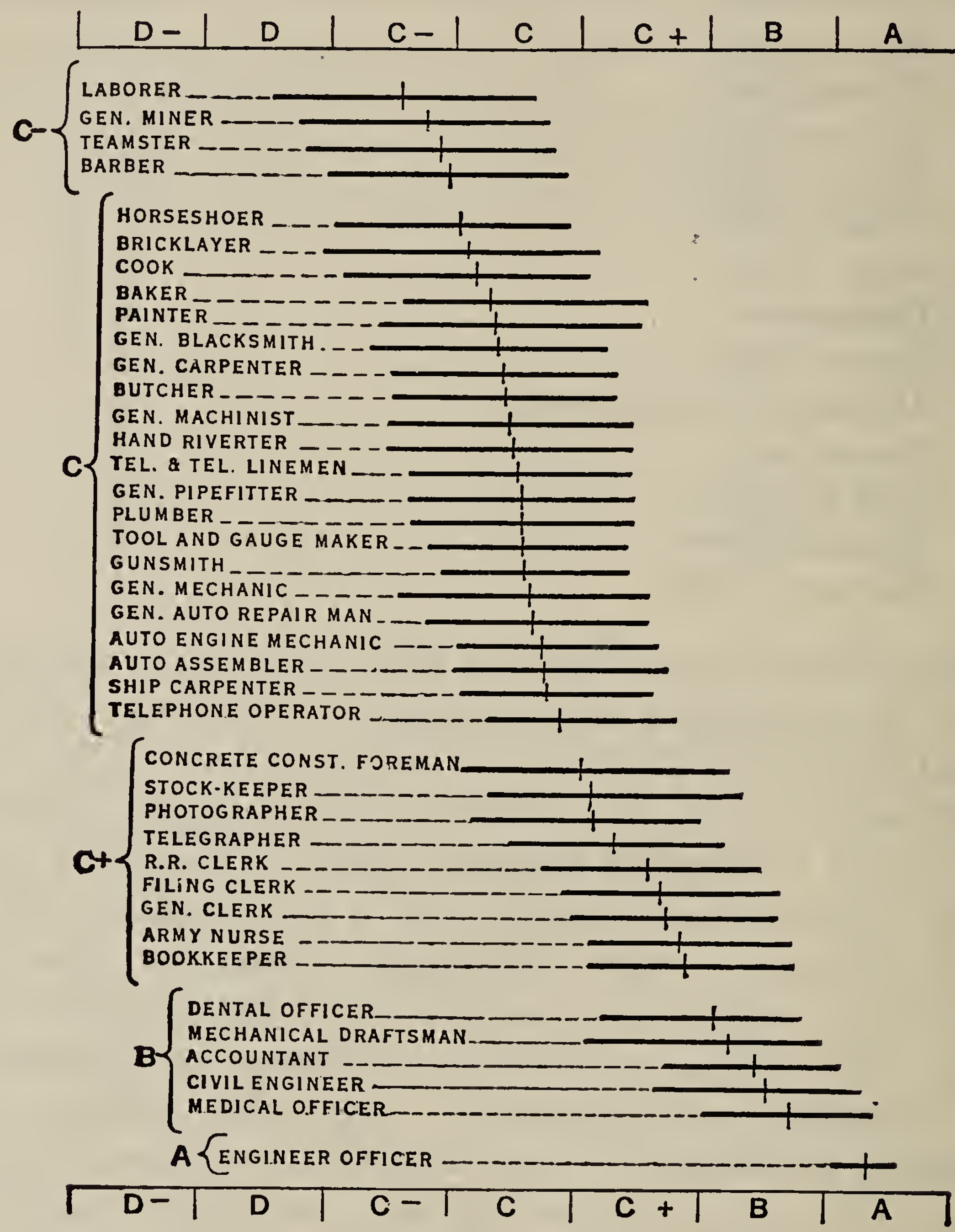


FIG. 27.—Occupational intelligence ratings. Letter grades on horizontal scale. Length of bar for each occupation is mid-range of 50 per cent; median point is shown by a cross line. (From *Memoirs of National Academy of Sciences,* Vol. XV.)

Taking the measurements as they stand, the 75 percentile unskilled laborer is up to the level of the median general mechanic, tool room expert, or automobile mechanic and up to the level of the 25 percentile mechanical engineer. The 75 per-

centile railroad clerk is at the level of the average accountant or civil engineer. The 75 percentile receiving or shipping clerk is at the level of the 25 percentile physician. This variability would be reduced by longer and repeated tests, but, unless the test as given has a very large probable error, it would still be enormous. It would still imply that there were in the occupations supposed to give little opportunity for the use of intellect, a very large number of gifted men and consequently a large unused surplus of intellect."

Again with reference to the army data, we must remember that they give us a picture of the situation as found in the army and this picture is not necessarily the same as exists in civil life. The discussion in Chapter 15 of the "Memoirs" (21) refers to the fact that the distribution of occupations in the army was influenced by the selective service act. "More than seven-tenths of those registering were placed in the deferred classes." It is obvious, therefore, that many occupations will not be adequately represented in the army occupational list. In some occupations the selective service act would tend to keep out of the army many of the most intelligent men. The farmers in the army were probably largely represented by farm laborers, because of the selective influence of the draft. Nevertheless, these army results are the best that we have at the present time and unquestionably indicate the importance of the factor of intelligence in various occupations. Further work of this sort may lead to the estimation of lower limits of intelligence for various occupations, in the sense that a certain minimum of intelligence is required for success in each occupation. An individual who does not possess the minimum required for a specific occupation should be advised not to plan to enter the occupation in question. Such limits of intelligence would be invaluable for the vocational counselor.

Fryer (22) has worked over the army data, amplified and corrected them, and constructed a table of intelligence scores for various occupational groups. Table 28 shows Fryer's results. The score range again shows the limits for the middle

TABLE 28

INTELLIGENCE LIMITS FOR VARIOUS OCCUPATIONS *

Intelligence Group	*Score Average*	*Score Range*	*Occupation*
A	161	110-183	Engineer (civil and mechanical)
	152	124-185	Clergyman
	137	103-155	Accountant
B	127	107-164	Physician
	122	97-148	Teacher (public schools)
	119	94-139	Chemist
	114	84-139	Draftsman
	111	99-163	Y.M.C.A. secretary
	110	80-128	Dentist
	109	81-137	Executive (minor)
C + ...	103	73-124	Stenographer and typist
	101	77-127	Bookkeeper
	99	78-126	Nurse
	96	74-121	Clerk (office)
	91	69-115	Clerk (railroad)
	86	59-107	Photographer
	85	57-110	Telegrapher and radio operator
	83	64-106	Conductor (railroad)
	82	57-108	Musician (band)
	81	59-106	Artist (sign letterer)
	81	60-106	Clerk (postal)
	81	57-109	Electrician
	80	62-114	Foreman (construction)
	80	56-105	Clerk (stock)
	78	54-102	Clerk (receiving and shipping)
	78	61-106	Druggist
	77	59-107	Foreman (factory)
	75	56-105	Graphotype operator

* From Fryer, D. "Occupational Intelligence Standards." By permission of *School and Society*.

Intelligence Group	*Score Average*	*Score Range*	*Occupation*
C	74	53-91	Engineer (locomotive)
	72	54-99	Farrier
	70	46-95	Telephone operator
	70	44-94	Stock checker
	69	49-93	Carpenter (ship)
	69	48-94	Handyman (general mechanic)
	69	46-90	Policeman and detective
	68	51-97	Auto assembler
	68	47-89	Engineman (marine)
	68	42-86	Riveter (hand)
	67	50-92	Toolmaker
	66	45-92	Auto engine mechanic
	66	45-91	Laundryman
	66	49-86	Gunsmith
	66	44-88	Plumber
	66	44-88	Pipefitter
	65	44-91	Lathe hand (production)
	65	43-91	Auto mechanic (general)
	65	43-91	Chauffeur
	65	42-89	Tailor
	65	44-88	Carpenter (bridge)
	64	43-88	Lineman
	63	40-89	Machinist (general)
	63	46-88	Motorcyclist
	63	41-86	Brakeman (railroad)
	62	31-94	Actor (vaudeville)
	61	40-85	Butcher
	61	44-84	Fireman (locomotive)
	61	39-82	Blacksmith (general)
	60	38-94	Shop mechanic (railroad)
	60	36-93	Printer
	60	40-84	Carpenter (general)
	59	40-87	Baker
	59	39-83	Mine drill runner
	59	38-81	Painter

Intelligence Group	*Score Average*	*Score Range*	*Occupation*
	58	37-85	Concrete worker
	58	40-83	Farmer
	58	37-83	Auto truck chauffeur
	58	37-82	Bricklayer
	57	41-81	Caterer
	57	39-71	Horse trainer
	56	38-76	Cobbler
	55	35-81	Engineman (stationary)
	55	34-78	Barber
	55	35-77	Horse hostler
	52	38-96	Sales-clerk
	52	33-74	Horseshoer
	51	31-79	Storekeeper (factory)
	51	26-77	Aeroplane worker
	51	31-74	Boilermaker
	50	33-75	Rigger
	50	30-72	Teamster
	49	40-71	Miner (general)
	48	21-89	Station agent (general)
C —	40	19-67	Hospital attendant
	40	19-60	Mason
	35	18-62	Lumberman
	35	19-57	Shoemaker
	32	16-59	Sailor
	31	20-62	Structural steel worker
	31	19-60	Canvas worker
	30	16-41	Leather worker
	27	19-63	Fireman (stationary)
	27	17-57	Cook
	26	18-60	Textile worker
	22	16-46	Sheet metal worker
	21	13-47	Laborer (construction)
D	20	15-51	Fisherman

fifty per cent. Only in a very general sense can such tables be used in vocational guidance.

Scott and Clothier (23) find a similar hierarchy of occupational levels for their mental alertness test given to various groups. Table 29 shows how the average score increases

TABLE 29

AVERAGE SCORE OF TYPICAL GROUPS *

Groups	*Average Score*
Sales force, employed to assist in holiday rush (women)	25
Sales force, department store (women)	27
Students, commercial business college (women)	28
Sales force, employed to assist in holiday rush (men)	29
Office boys	31
Sales force, department store (men)	33
Machine operators (men)	33
Job foremen	38
All office employees (women)	40
Foremen	41
Nurses	42
Rotary Club members	46
Executives of progressive firm	51
Supervisors (manufacturing plant)	52
Sales executives	54
Students (college of arts and sciences)	54
Students (medical school)	56
Engineering students	57
College presidents (small colleges)	58
Students (succeeding in examinations for internship)	60

from the sales people up to office employees and then up to executives.

It is interesting to note that the results of testing 6,000 Japanese soldiers by Awaji (28) showed the following ranking of occupations: (1) professional; (2) clerical; (3) business;

* From Scott, W. D., and Clothier, R. C. Personnel Management. By permission of McGraw-Hill Book Company, publishers.

(4) skilled; (5) semi-skilled; (6) agriculture; (7) labor. The author regards his group as a random sampling of Japanese adult men of ages 20 to 21, because of compulsory military service.

Selection and Guidance.—By far the most common use and, perhaps, the most important use of intelligence tests in industry and business has been as an aid in the selection of individuals for various kinds of work. As a rule the intelligence test is only a part of the means employed for selection. It is very obvious that it must always remain a part. The judging and classification of men is a difficult and complex task, as Thorndike (18) and Kelley (19) have shown.

The general value of intelligence tests in industry has been pointed out by many writers (e.g., Scott, 16; Whipple, 16) and Lamb (19) has given an account of the use of such tests in a large manufacturing plant, particularly as an aid to the employment manager.

A common method of selecting and evaluating tests is described by Hollingworth (16) as follows: "In a recent investigation an attempt was made to discover a set of mental tests which would aid in the selection of efficient workers in a specific field. Thirty workers who were already employed under fairly comparable conditions of work were taken as subjects in a preliminary search for tests of value. These thirty people were each put through a series of association tests, of the familiar laboratory form, naming opposites, naming colors and forms, completing mutilated passages, following hard directions, giving responses bearing specified relations to stimulus words, cancellation, number checking, etc. While these tests were in progress, during a period of several days, the thirty workers were rated by three supervisors who were familiar with their work at the actual task, and who had for some time been observing their performance with a view to making subsequent judgments. Each superior arranged the thirty workers in an order of merit, according to his or her impression of their relative efficiency. The judgments of the three supervisors

were then averaged and each worker assigned a final position on the basis of these averages. This was believed to be as accurate a measure of actual ability as could be secured under the complex conditions of work. The results of these ratings were then compared with the results of the mental tests. Some of the tests were found not to correlate with the ratings for actual working efficiency. Three tests showed definite and positive correlations. When results from these three tests were combined, the records correlated with the ratings by a coefficient of fifty-five per cent. These three tests were then accepted as having value in the selection of good operators, and search was continued for further tests which might also yield positive correlations."

For guidance purposes Darsie (25) suggests that the I.Q. may indicate types of education and types of vocation. Further education would be suggested by the I.Q. as follows:

Above 115	College
100-114	High school
80-99	Vocational high school
70-79	Industrial junior high school
Below 70	Special school

His occupations are grouped into three large classes and the I.Q. indications are:

I.Q. 120 and over	professional
I.Q. 119 to 90	clerical and skilled work
I.Q. 90 and below	semi-skilled work

Burt (23-24) also gives an ideal distribution according to I.Q. for guidance purposes:

I.Q.	*Probable Per Cent*	*Probable Vocation*
Above 150	0.1	Professions
130 to 150	1 to 2	Clerks, teachers
115 to 130	10	Clerks, skilled workers
100 to 115	38 to 39	Trade, skilled labor

I.Q.	*Probable Per Cent*	*Probable Vocation*
85 to 100	40	Semi-skilled
70 to 85	10	Unskilled
50 to 70	1	Casual, domestic labor
Below 50	0.1	Institutional cases

These standards of Burt are very much higher than those of Darsie.

Gaw (26) describes an actual attempt to give vocational advice to 100 children just ready to leave school. The advice was based upon the results of abstract and concrete intelligence tests as well as upon all other available information about each case. Two years later many of these cases were followed up. Most of those who had followed the vocational recommendations were satisfied with their work, but of those who had gone into a different type of work from that recommended only about half were satisfied with their jobs.

In the U. S. Civil Service intelligence tests of various kinds are used as part of the examinations given for various positions and are considered as valuable aids in the selection of applicants (O'Rourke).

Poull (22) shows that different types of intelligence tests may have value as indicators of different sorts of interest among children. She found that children who expressed interest in mechanical occupations obtained higher scores on the Pintner Non-language Test than on the National Verbal Intelligence Test. In this way different kinds of intelligence tests may be of significance in vocational guidance.

We may now briefly survey the different occupations in which intelligence tests have been used or in which tests have been employed to help in the selection of applicants for such occupations.

Clerical Workers and Executives.—By far the largest use of intelligence tests has been made in connection with such positions. Thorndike and Scott have for a long time made use of intelligence tests in the examination of applicants for clerical

and executive positions. Notable in this connection are Thorndike's tests for the Metropolitan Life Insurance Company. Ordinarily in the construction of tests for selection purposes, the intelligence test is only one part of the total examination. Thus Thurstone (19) and Carney (19) use intelligence tests as part of their general clerical examination, while Bregman (21) uses the conventional intelligence tests almost entirely. Flanders (18) makes use of the Stanford-Binet Scale. Gardener (17), giving a report of the practical situation in the employment manager's office, includes intelligence tests as one of five valuable sources of data about the applicant, the other four being personal interview, physical examination, interview with foreman and references from other employers. Most of these reports stress the importance and value of intelligence tests as a part of the total sizing-up of the individual. Intelligence alone, however, is not enough and cannot be the sole criterion. Flanders found no positive relation between degree of intelligence and the rating of the men as to all-round efficiency, dependability, loyalty and coöperativeness. At any particular job we may find men possessing a much greater amount of intelligence than the job in question demands. They may be too intelligent for the job. Unless they are well endowed with other useful moral and social qualities, their attitude towards their work may be less desirable, and hence their total value to the firm of less account than that of other less intelligent individuals.

Bills' (23) study is interesting in this connection. An intelligence test was given to 133 clerical workers. The correlation between intelligence and difficulty of the job was + .22. Two and a half years later the correlation was + .41 for those still at work. It was found that those scoring low tended to leave, but also that those with high intelligence scores had left the low grade jobs.

An intelligence test was given by Scudder (29) to 262 veterans in a rehabilitation course in bookkeeping and accounting. The median score for 170 who finished the course was 142; for

92 who dropped out it was 112. There were 103 of the 142 in employment about 3 or 4 years later, and those scoring high on the test had done much better than those scoring low. The average monthly increase in wages was for the

Upper 10 = $145
Middle 10 = $73
Lowest 10 = $41

From these results the author suggests minimum scores on the Terman Mental Ability Test for those planning to become bookkeepers and accountants.

When we come to deal with a select group of higher executives, Bingham and Davis (24) have shown that an intelligence test of the army type does not differentiate as to business success. For 102 individuals all scoring above the army median score, the correlation between intelligence score and business success was −.10. They conclude that "superiority in intelligence, above a certain minimum, contributes relatively less to business success than does superiority in several non-intellectual traits of personality."

Salesmen.—The use of tests for the selection of salesmen has been stimulated most of all by the work of Scott. His tests and modifications of his tests have been used by many stores and industrial concerns. Hollingworth and Poffenberger (17) show how intelligence tests correlate with the average salary per year of experience of a number of salesmen engaged in selling all manner of commodities. Oschrin (18), and later Bregman (21), report the use of tests for the selection of saleswomen in a New York department store. Various tests and combinations of tests were used and they seemed to discriminate between good and poor sales-clerks. The tests were then used for the selection of new employees. The most satisfactory tests proved to be Directions, Judgment and Comparison—all essentially intelligence tests.

Anderson (29) discusses the uses of intelligence tests in a large department store, finds a slight correlation with sales

ability and gives the following percentage distribution of the I.Q.s of 500 sales-clerks tested by the Otis S.A. Test:

I.Q.	*Per Cent*
Above 110	5.4
90-109	40.6
80-90	34.4
70-80	15.6
60-70	4.0
Below 60	0.0

From this it can be seen that there are sales-clerks of all degrees of intelligence, except feeblemindedness. Fifty per cent lie between 70 and 90, in the dull zone of abstract intelligence.

Stenographers.—Tests for the selection of stenographers, typists and comptometer operators have been reported by Rogers (17) and Bills (21). Both authors feel that intelligence tests are valuable aids in such selection. Bills also introduces the Downey Will-Temperament Tests. She finds that the general intelligence test is the most efficient for eliminating failures. "Failures can be predicted by the tests with over 85 per cent accuracy." The mean score on the intelligence test for various groups of stenographers is "for secretaries 144; for stenographers rated 'good,' 110; for stenographers rated as 'getting by,' 65; and for those failing, 63." A certain amount of general intelligence is necessary for success in these occupations and persons not coming up to a certain standard should be advised against training for the work.

Nurses.—Earle (26) reports the results of giving the Army Alpha to 212 student nurses with a mean score of 113, and a median of 110. This median score is a little higher than the median reported in the army results. She found no correlation between intelligence and the ratings of supervisors on the characteristics necessary for an ideal nurse. Young (24) also reports intelligence test results for 122 student nurses. These test eight percentile points above high school seniors, and seven percentile points above teachers in training. This author also

found that the student nurses "rating highest on the intelligence test, made distinctly the highest average in both their theoretical studies, and their efficiency record on practical work."

Telegraphers.—Jones (17) and Thurstone (19) report investigations with mental tests for selecting telegraphers. Jones finds a high correlation between ability to succeed in a school for telegraphers and six tests, whereas Thurstone does not find such a high correlation. Just how large a part general intelligence plays in success in telegraphy is, therefore, doubtful. Thurstone feels that "general intelligence tests are not as valuable for diagnosing ability to learn telegraphy as for measuring general intelligence. Ability in telegraphy is probably a special ability."

Policemen and Firemen.—As part of a civil service examination for the selection of policemen and firemen Terman (17) reports the use of the short form of the Stanford-Binet along with some educational tests. The range in mental ages was from 10 to 18, with a median mental age of 13.5 or an I.Q. of 84. It was recommended that all men with I.Q.s below 80 be rejected, and this recommendation was accepted. There is no record of the future success of the men who were selected, nor can we know whether those rejected might not have developed into good policemen or firemen.

There are three reports of Army Alpha given to policemen:

Author	*Average Score*	*n*	
Anon (26)	76	64	Applicants
Fernald and Sullivan (26)	82	1,712	Los Angeles
Merrill (26-27)	104	113	Palo Alto

These scores are above the average for the army draft. A minimum score of 65 is suggested by the civil service report (Anon, 26). The Merrill report shows no difference in score between the men who were discharged and the men who remained. It does show a great difference in average score for four men who

left for better jobs, their average score being 171. Another report by Hamill (23) gives results for 50 policemen on the Pintner Non-Language, and according to this test the percentage distribution was: 4% inferior; 58% average; 38% superior.

Semi-skilled and Unskilled Workers.—The lower down the scale of industry we go, the less valuable do our present intelligence tests appear to be for the selection of workers. Otis (20) finds no relation between a performance intelligence examination and the productive ability of 400 workers in a large silk manufacturing concern. A large percentage of these mill workers were foreign or illiterate and so the performance type of examination was used. Otis comes to the conclusion that "intelligence is not only not required in a modern silk mill for most operations, but may even be a detriment for steady efficient routine work. What qualities are required remain to be sought. Whether they are measurable is doubtful. They may be stolidity, patience, inertia of attention, regularity of habit, etc."

With a group of eight tests, mostly of the sensory-motor type, Link (18) finds in general positive correlations with efficiency in shell inspecting and shell gauging in an ammunition factory. He concludes that three of these tests will prove useful in the selection of applicants for these jobs. His work bears out the conclusions of Otis in the sense that the only test of the eight used by Link that might be called a general intelligence test correlates very slightly with the abilities in question and is not retained by him for future use for selection purposes.

Similarly Viteles (25) finds no correlation between intelligence and ratings of the efficiency of motormen.

Aviation.—Although the aviator can hardly be called an employee in the sense of that term as used in this chapter, mention must be made of the use of tests for the selection of aviators. We have already quite an extensive literature on the subject and Dockeray and Isaacs (21) have given us an excellent summary of psychological research work carried on in

Italy, France, England and the American Expeditionary Forces. Reaction time tests, tests for emotional stability, for equilibrium, steadiness and the like are common. Very extensive were the investigations upon the effect of altitude on various types of response. Stratton and others (20) report tests of judgment of curves, of relative speeds, reaction times, muscular exertion and fatigue. And Henmon (19) gives correlations of flying ability with mental alertness, emotional stability, swaying, visual reaction and several other tests. It is interesting to note that mental alertness and emotional stability show the highest correlations with flying ability of the ten tests reported.

Thorndike's Mental Alertness Test was given to several hundred cadets at the aviation ground schools and showed a correlation of 50 with their average grade in school work. Thorndike recommends the use of this intelligence test as one test among nine for the selection of prospective aviators. Other factors of importance are athletic achievement, athletic mechanical ability, emotional stability and a delicate sense of balance. Achievement as an aviator, Thorndike (19-20) shows, correlates as follows:

With education	About zero
With athletic ability	About .15
With athletic-mechanical ability ..	About .2 to .3
With wage in civil life	About .1
With social achievement	About .04

Dockery and Isaacs (21) stress the factor of intelligence when they say, "What seems most needed by the aviator is intelligence, that is, the power of quick adjustment to a new situation and good judgment. He need not be so quick in motor adjustments, provided he thinks clearly or makes quick mental adjustments."

The Unemployed.—No results comparable to the army results can be given for the unemployed. In the investigations to be mentioned the tests employed are so different as to afford no basis for comparison. Only in a general way can we say

that the unemployed as a group give the impression of being very low in intelligence. Pintner and Toops (17) tested 94 cases in one city and 40 in another, and give the following percentage distribution:

	City A	*City B*
Feebleminded	28.7	7.5
Borderline	29.8	25.0
Backward	28.7	32.5
Normal	8.5	20.0
Bright	4.3	15.0

Johnson (17) tested 107 unemployed and estimates 21 per cent of them as feebleminded. No importance should be attached to the actual percentage of feeblemindedness obtained by these workers. Such results should be taken merely as an indication that low intelligence is likely to be very common among the unemployed. The lower the intelligence of the individual, the harder will it be for him to obtain a job and the harder to retain a job, because he will in general be the one to be dismissed when times are hard and jobs are few.

Intelligence and Complexity of Task.—Certain workers have endeavored to establish minimum mental levels necessary for success in certain jobs. Thus Burr (24) gives the following minimum levels in millinery and other simple occupations:

M.A.	
7 to 8	Packing small articles not easily damaged, e.g., powder puffs
8 to 9	Certain operations in pencil making
9 to 10	Crude hand sewing; cutting and pasting in paper-box making; sewing hat linings; packing difficult articles, e.g., hair nets
10 to 11	Stock-keeping; winding cotton and wool braid
11 to 12	Covering buckram foundations, sewing in wire edge and facings
12 to 13	Machine operating for straw millinery, window shades, garters, etc.
13 to 14	Assembling of parts requiring judgment; completion of entire garment

Cowdery (22) and Vanuxem (25) have worked out similar tables for tasks in institutional life.

Summary.—In this chapter we have attempted to describe the work done in intelligence testing in so far as it refers to the man engaged in some occupation or trade, and we can divide this work into two parts. In the first place we have the measurement of the intelligence of men in different occupations, and in the second place, the search for tests which will predict success in some occupation.

The facts as to the intelligence of employees give us a hierarchy of occupations arranged according to intelligence. These facts show how intelligence is distributed at the present time in various lines of work. A natural, but probably very rough, selective process is going on all the time, whereby the less intelligent are relegated to simpler occupations and only the more intelligent survive in more complex occupations. The rough facts as we know them to-day may serve, therefore, as a tentative guide for vocational advice.

The more immediately practical work is the use of tests for selection. Here we find that intelligence tests are useful, but that they alone will not suffice for accurate prediction. To fill successfully a position and to hold it requires other attributes besides general intelligence, although a certain minimum of intelligence will be necessary for each particular job. Hence the use of intelligence tests for selection purposes will vary greatly according to the job to be filled. Their use, furthermore, will depend upon the immediate aim of the selection, whether it be to fill just one particular job or to select a candidate who can fill the job in question, but in addition grow and develop into more important positions in the firm.

Obviously there is a close relationship between the first type of work, the determination of the intelligence level of various occupations, and the second type of work, the selection of candidates for jobs. Unquestionably the distribution of intelligence among different occupations to-day is not the most economical nor the most effective. There must be thousands of

workers who are forced to toil at tasks far below their intelligence level, and on the other hand hundreds who are struggling to fill positions which are beyond their intellectual reach. Both are maladjusted, and maladjustment means unhappiness and inefficiency. We may hope, therefore, that in the future intelligence tests and psychological tests of all sorts will prove a means whereby a scientific adjustment of the worker to his work may be attained, resulting in increased satisfaction to the worker and increased efficiency in the work.

REFERENCES

Anderson, V. V. (29). *Psychiatry in Industry.* New York.

Anon. (26). "Partially Standardized Tests for Patrolmen." *Public Personnel Studies,* 4, 122-151.

Awaji, Y. (28). "Intelligenzprüfung im Japanischen Heere." *Zeitsch. f. angew. Psychol.,* 30, 81-118.

Ayres, L. P. (13). "Psychological Tests in Vocational Guidance." *J. of Ed. Psych.,* 4, 232-237.

Bills, M. A. (21). "Methods for the Selection of Comptometer Operators and Stenographers." *J. of Appl. Psych.,* 5, 275-283.

Bills, M. A. (21). "A Test for Use in the Selection of Stenographers." *J. of Appl. Psych.,* 5, 373-377.

Bills, M. A. (23). "Relation of Mental Alertness Test Score to Positions and Permanency in Company." *J. of Appl. Psych.,* 7, 154-156.

Bingham, W. V. and Freyd, M. (26). *Procedures in Employment Psychology.* Chicago.

Bingham, W. V. and Davis, W. T. (24). "Intelligence Test Scores and Business Success." *J. of Appl. Psych.,* 8, 1-22.

Bregman, E. O. (21). "A Study in Industrial Psychology—Tests for Special Abilities," *J. of Appl. Psych.,* 5, 127-151.

Burr, E. T. (24). "Minimum Intellectual Levels of Accomplishment in Industry." *J. of Person. Res.,* 3, 207-212.

Burt, C. (23-24). "The Principles of Vocational Guidance." *Brit. J. of Psych.,* 14, 336-352.

Burtt, H. E. (26). *Principles of Employment Psychology.* Boston.

Carney, C. S. (19). "Some Experiments with Mental Tests as an Aid in the Selection and Placement of Clerical Workers in a Large Factory." *Bulletin of the Extension Division,* Indiana University, Vol. V, No. 1, 60-74.

Chapman, J. C. (21). Trade Tests. *The Scientific Measurement of Trade Proficiency.* New York.

Cowdery, K. M. (22). "Measures of General Intelligence as Indices of Success in Trade Learning." *J. of Appl. Psych.,* 6, 311-330.

Darsie, M. L. (25). "A Method of Reporting the Significance of Intelligence Tests to Parents and Teachers." *Sch. and Soc.*, 22, 597-600.

Dockeray, F. C. and Isaacs, S. (21). "Psychological Research in Aviation in Italy, France, England and the American Expeditionary Forces." *J. of Comp. Psych.*, 1, 115-148.

Earle, M. G. (26). "The Relation between Personality and Character Traits and Intelligence." *J. of Appl. Psych.*, 10, 453-461.

Fernald, G. M. and Sullivan, E. B. (26). "Personnel Work with the Los Angeles Police Department." *J. of Delinq.*, 10, 252-267.

Flanders, J. K. F. (18). "Mental Tests of a Group of Employed Men Showing Correlations with Estimates Furnished by Employer." *J. of Appl. Psych.*, 2, 197-206.

Freyd, M. (26-27). "Selection of Typists and Stenographers. Information on Available Tests." *J. of Person. Res.*, 5, 490-510.

Fryer, D. (22). "Occupational Intelligence Standards." *Sch. and Soc.*, 16, 273 ff.

Gardener, H. L. (17). "The Selection Problem of Cheney Bros." U. S. Bureau of Labor Statistics. *Bulletin 227*, 120-125.

Gaw, F. et al. (26). "A Study in Vocational Guidance." Indus. Fatigue Res. Board, *Report No. 33*, London, H. M. Stationery Office.

Hamill, G. (23). "The Application of the Pintner Group Test to Misdemeanants." *J. of Delinq.*, 8, 158-168.

Henmon, V. A. C. (19). "Air Service Tests of Aptitude for Flying." *J. of Appl. Psych.*, 3, 103-109.

Hollingworth, H. L. (16). *Vocational Psychology.* New York.

Hollingworth, H. L. and Poffenberger, A. T. (17). *Applied Psychology.* New York.

Johnson, G. R. (17). "Unemployment and Feeblemindedness." *J. of Delinq.*, 2, 59-73.

Jones, E. S. (17). "The Woolley Test Series Applied to the Detection of Ability in Telegraphy." *J. of Ed. Psych.*, 8, 27-34.

Kelley, T. L. (19). "Principles Underlying the Classification of Men." *J. of Appl. Psych.*, 3, 50-67.

Kitson, H. D. (25). *The Psychology of Vocational Adjustment.* Philadelphia.

Lamb, J. P. (19). "Intelligence Tests in Industry." *Ind. Man.*, 58, 21-23.

Link, H. C. (18). "An Experiment in Employment Psychology." *Psych. Rev.*, 25, 116-127.

Link, H. C. (19). *Employment Psychology.* New York.

Memoirs (21). The National Academy of Sciences, Vol. XV, Part III, Chap. 15.

Merrill, M. A. (26-27). "Intelligence of Policemen." *J. of Person. Res.*, 5, 511-515.

O'Rourke, L. J. See *Public Personnel Studies.* Bureau of Public Personnel Administration, Washington, D. C.

Oschrin, E. (18). "Vocational Tests for Retail Saleswomen." *J. of Appl. Psych.*, 2, 148-155.

Otis, A. S. (20). "The Selection of Mill Workers by Mental Tests." *J. of Appl. Psych.,* 4, 339-341.
Personnel System of the United States Army (19). Vols. I and II, Washington, D. C.
Pintner, R., and Toops, H. A. (17-18). "Mental Tests of Unemployed Men." *J. of Appl. Psych.,* 1, 325-341, and 2, 15-25.
Poffenberger, A. T. (27). *Applied Psychology, Its Principles and Methods.* New York.
Poull, L. E. (22). "Interests in Relation to Intelligence." *Ungraded,* Vol. VII, Nos. 7, 8 and 9.
Rogers, H. W. (17). "Psychological Tests for Stenographers and Typewriters." *J. of Appl. Psych.,* 1, 268-274.
Rossy, C. S. (17). "Mental Examinations for Employees." *Ind. Man.,* 54, 359-365.
Scott, W. D. (16). "Selection of Employees by Means of Quantitative Determinations." *Annals of Am. Acad. of Pol. and Soc. Sci.,* 65, 82-193.
Scott, W. D., and Clothier, R. C. (23). *Personnel Management.* Chicago.
Scudder, K. J. (29). "The Predictive Value of General Intelligence Tests in the Selection of Junior Accountants and Bookkeepers." *J. of Appl. Psych.,* 13, 1-8.
Stratton, O. M. et al. (20). "Psychological Tests for Selecting Aviators." *J. of Exp. Psych.,* 3, 405-423.
Terman, L. M. (17). "A Trial of Mental and Pedagogical Tests in a Civil Service Examination for Policemen and Firemen." *J. of Appl. Psych.,* 1, 17-29.
Thorndike, E. L. (18). "Fundamental Theorems in Judging Men." *J. of Appl. Psych.,* 2, 67-76.
Thorndike, E. L. (19). "Scientific Personnel Work in the Army." *Science,* 49, 53-61.
Thorndike, E. L. (19-20). *The Selection of Military Aviators.* U. S. Air Service, Vol. I, No. 5; Vol. II, Nos. 1 and 6.
Thurstone, L. L. (19). "A Standard Test for Office Clerks." *J. of Appl. Psych.,* 3, 248-251.
Thurstone, L. L. (19). "Mental Tests for Prospective Telegraphers." *J. of Appl. Psych.,* 3, 110-117.
Toops, H. A. (21). *Trade Tests in Education.* T. C. Contribs. to Educ., No. 115. Columbia University.
Vanuxem, M. (25). *Education of Feebleminded Women.* T. C. Contribs. to Educ., No. 174, Columbia University.
Viteles, M. S. (21). "Tests in Industry." *J. of Appl. Psych.,* 5, 57-63.
Viteles, M. S. (25). "Research in Selection of Motormen." *J. of Person. Res.,* 4, 100-115 and 173-199.
Whipple, G. M. (16). "The Use of Mental Tests in Vocational Guidance." *Annals of Am. Acad. of Pol. and Soc. Science,* 65, 193-204.
Young, H. H. (24). "Intelligence Ratings and Success of Nurses in Training." *J. of Appl. Psych.,* 8, 377-389.

CHAPTER XXIII

THE SEXES

The General Problem.—The literature on the differences between men and women is enormous. The subject seems of perennial interest. The centuries-old popular and universal belief in the inferiority of women, both physically and mentally, has gradually been broken down in recent times. The vast differences in mental traits have shrunk to very small differences or entirely disappeared. At the same time the hypothesis of the greater variability of the male sex has been maintained in order to account for the seemingly greater number of eminent men than women. But this hypothesis has been challenged and data presented to confute it, and the greater number of eminent men has been explained on the basis of greater opportunity for male eminence and restriction of eminence very largely to masculine walks of life. This chapter will deal with a very small fraction of the total problem of sex differences, namely a comparison of the sexes on intelligence tests. We shall present a sampling of the data available and inquire into the difference between the sexes and the comparative variability of the sexes.

Test Results.—A survey of all the results of intelligence tests shows very few reports giving separate findings for the two sexes. Most workers have taken for granted that there are no real differences in intelligence between boys and girls. In the establishment of age or grade norms for intelligence tests, most workers have massed boys and girls together, tacitly assuming that sex makes no difference. We assign M.A. from score without regard to sex.

Table 30 gives a summary of the chief reports of group in-

TABLE 30

Sex Differences on Intelligence Tests

Author	*Test*	*n*	*Central Tendency*	*Variability*
Pressey (18)	Pressey	2,544	Girls superior	Boys more variable
Pintner (24)	Non-Language	2,270	No difference	
Colvin and Macphail (24)	Brown	3,333 H. S.	Boys superior	
St. Louis Report (25)	N. I. T.	10,138 ages 7–16	Girls superior	
Thorndike (26)	I. E. R.	2,500 ages 13–18	Boys superior	
Thorndike (26)	I. E. R.	11,600 ages 13–18		Boys more variable
Whipple (27)	N. I. T.	2,198 age 11	Girls superior	
	Illinois	1,919 ages 8–13	Girls superior	
Whipple (27)	Army Alpha	834 H. S.	Boys superior	
Winsor (27)	Army Beta	2,321 age 10	Test 2—Boys superior	Girls more variable
			Test 5—Girls superior	Boys more variable
Brigham (26–29)	N. I. T.	389 ages 10–11	Girls superior	Boys more variable
	C. E. B.	Various	Girls superior (except in one report)	Boys more variable
Lincoln (27)	Dearborn	3,422 ages 7–16	No real difference	Boys more variable
Book and Meadows (28)	Indiana	5,925 H. S. Seniors	Boys superior	Boys more variable
Commins (28)	Multimental	100 per age and sex: ages 9–14	Girls superior	
Broom (30)	Terman	1,200 H. S.	No difference	No real difference

telligence tests where the difference between the sexes has been made a special object of study. If we examine this table, we find that there are 15 reports in which the differences between boys and girls in central tendency are noted. Winsor reports on two tests and one of the Thorndike reports deals only with variability. The girls are superior in seven, the boys in four and no difference is reported in four. The four reports showing the boys superior are all high school reports and several of our authors point out that the slight superiority of the boys is probably due to the narrower selection of boys in high school, the less able dropping out to go to work more frequently than the less able girls. The seven reports showing the girls superior are mainly concerned with the elementary school ages. Suggested reasons for this superiority are the earlier maturity of girls, the greater docility of girls in the educational régime and hence their higher scores on those tests influenced by educational achievement. No real difference between boys and girls is found on the Dearborn Test with many cases ranging in age from 7 to 16, on the Pintner Non-Language Test at ages 10 and 12, on the Beta Test as reported by Winsor, or on the Terman Group Test with high school children. Even where differences in central tendency are reported they are frequently small. As almost every author points out, any difference which may exist is small, and the similarity rather than the difference is the striking phenomenon. Small differences between groups tested may easily occur owing to differences in the selection of samples. But, where the sampling is probably most similar for both sexes, namely, in the elementary school ages, we do find the preponderance of evidence in favor of the girls. The reports which find no difference at these ages are those of Pintner (24) and Lincoln (27), and it is to be noted that these authors are dealing with tests that have much less verbal material than the other tests in which the girls are superior. These results, therefore, fit roughly into the general picture of the slight superiority of the female in verbal material (Good-

enough, 27), whether due to early conditioning or to an inherent difference between the sexes in mental traits.

Again, we repeat, the difference is so slight that for educational purposes it may be entirely disregarded. For intelligence testing purposes with our present rough measures, the difference between the sexes may also be disregarded, but in more accurate work in the future we may have to introduce more non-verbal material to counteract the special ability of girls in verbal material, should it be conclusively proved that they do possess such a special ability.

Variability.—As regards variability in intelligence tests, we find from our table that nine reports discuss this problem. Seven of these find the boys more variable, one the girls and one (Winsor on Army Beta) that the variability shifts from sub-test to sub-test. In general, therefore, the results of intelligence testing seem to favor the theory of a slightly greater variability in the male sex.

High and Low I.Q.s.—To explain the supposedly greater eminence of men, the hypothesis of the greater variability of the male sex has been suggested. In addition to this it has been pointed out that a greater number of extreme male variants at both ends of the distribution would readily account for male superiority. With reference to general intelligence, then, let us enquire whether we find a greater number of boys with very high and very low I.Q.s.

Terman's data (Terman et al. 25) give us the best answer with reference to very high I.Q.s. In his search for children with I.Q.s above 130, he was very careful to avoid any sex preference. Nevertheless, his main experimental group of 643 consisted of 352 boys and 291 girls, that is, 54.7 per cent boys and 45.3 per cent girls, or a ratio of 121 boys to 100 girls. An addition of 33 later cases reduced this ratio to 116 boys to 100 girls. Another of his groups has a ratio of 123.9 boys to 100 girls, and his high school group a ratio of 212.3 boys to 100 girls. The well-known excess of male to female births is about

105 or 106 boys to 100 girls. But allowing for this excess we still find a higher ratio for boys in the gifted group.

Terman gives the percentages of each sex for the extremely high I.Q.s as follows:

I.Q.	*160 up*	*170 up*	*180 up*	*190 up*
Boys	18.5	6.2	2.6	0.0
Girls	16.5	7.2	2.0	1.0

and says that "except for the non-excess of boys in the highest I.Q. range, the facts we have presented are in harmony with the hypothesis that exceptionally superior intelligence occurs with greater frequency among boys than girls." Terman discusses thoroughly the possibility of biased selection, the sex ratio in the families studied, and the differential death rate of embryos, and he concludes, "the true cause of the sex ratio cannot be determined from our data. It may be either variability or the differential death rate of embryos."

Let us now turn to the other end of the distribution and enquire whether we find a greater number of males of very low I.Q. The British Mental Deficiency Committee (Report, 29) probably represents the most careful survey of a large area. It reports the incidence of idiots 30 per cent higher for boys than for girls. Wallin (30), reporting on 1019 cases referred to a psychological clinic, found that 1.0 per cent of the boys and 0.9 per cent of the girls had I.Q.s below 35. With another group of 1,114 cases, he found that 1.6 per cent of the boys and 0.9 per cent of the girls had I.Q.s below 35. In Bridgman's (29) report of 3,675 clinic cases, we find 1.5 per cent males and 1.2 per cent females diagnosed as idiots. But, Hollingworth (22) found only 6.4 per cent of the boys and 7.8 per cent of the girls with I.Q.s below 35 among 769 cases presented for examination. She presents data to show that fewer girls are segregated than boys and that the girls who are segregated escape longer than do the boys. She points out the impossibility of accepting data from institutions as proving the greater number of male feebleminded. The only evidence of

a preponderance of male idiots rests then upon such a survey as carried out by the British Mental Deficiency Committee, and it fits into the general theory of greater male variability in general intelligence with a preponderance of males at both extremes of the distribution.

Special Tests.—There are a great number of other reports on sex differences dealing with all sorts of tests. As they are not all general intelligence tests, we can only deal briefly with some of them. Pintner and Anderson (17) found no sex differences in the Healy Picture Completion Test. McFarlane (25) found boys in general superior to girls in tests of practical intelligence. In number or arithmetical tests boys are frequently found to be superior, whereas in language tests girls are generally superior, as also in tests of memory. It would seem, therefore, that within the range of those types of material usually used in intelligence tests, there may be sex differences which disappear in the composite known as general intelligence. Whether these differences are due to early conditioning or whether they are native has not been determined.

Conclusions.—This brief survey of sex differences in general intelligence leads us to the conclusion that no real difference has been as yet determined. If there is one, it may be a slight superiority of girls in verbal intelligence tests. Boys would seem to be slightly more variable than girls, and boys seem to be somewhat more numerous at the extreme ends of the distribution of intelligence. All of these differences in central tendency and variability appear to be extremely small. Differences in special traits may be present and remain to be more thoroughly investigated. For educational purposes, the differences in central tendency and variability are so small as to be of no significance.

REFERENCES

Book, W. F., and Meadows, J. L. (28). "Sex Differences in 5,925 High School Seniors in Ten Psychological Tests." *J. of Appl. Psych.*, 12, 56-81.

Bridgman, O. (29). "The Sex of Mentally Deficient Individuals." *Ment. Hyg.*, 13, 62-69.

Brigham, C. C. (26-29). *Annual Reports on Scholastic Aptitude Tests.* College Entrance Board.

Broom, M. E. (30). "Sex Differences in Mental Ability Among Junior High School Pupils." *J. of Appl. Psych.*, 14, 83-90.

Colvin, S. S., and MacPhail, A. H. (24). "Intelligence of Seniors in the High Schools of Massachusetts." *U. S. Bureau of Educ. Bull.*, No. 9.

Commins, W. D. (28). "More About Sex Differences." *Sch. and Soc.*, 28, 599-600.

Ellis, R. S. (28). *Psychology of Individual Differences.* New York.

Goodenough, F. L. (27). "The Consistency of Sex Differences in Mental Traits at Various Ages." *Psych. Rev.*, 34, 440-462.

Henmon, V. A. C., and Livingston, W. F. (22). "Comparative Variability at Different Ages." *J. of Ed. Psych.*, 13, 17-29.

Hollingworth, L. S. (22). "Differential Action upon the Sexes Which Tend to Segregate the Feebleminded." *J. of Ab. and Soc. Psych.*, 35-57.

Lincoln, E. A. (27). *Sex Differences in the Growth of American School Children.* Baltimore.

Louttit, C. M. (25-26). "A Bibliography of Sex Differences in Mental Traits." *Training School Bull.*, 22, 129-138.

McFarlane, M. (25). "A Study of Practical Ability." *Brit. J. of Psych.*, Mono. Suppl., No. 8.

Pintner, R. (24). "Results Obtained with the Non-Language Group Test." *J. of Ed. Psych.*, 15, 473-483.

Pintner, R., and Anderson, M. M. (17). *The Picture Completion Test.* Baltimore.

Pressey, L. W. (18). "Sex Differences Shown by 2,544 School Children." *J. of Appl. Psych.*, 2, 323-340.

Pyle, W. H. (24). "Sex Differences and Sex Variability in Learning Capacity." *Sch. and Soc.*, 19, 352.

Pyle, W. H. (25). "The Relation of Sex Differences to the Kind of Material Used." *J. of Ed. Psych.*, 16, 261-264.

Report of the Mental Deficiency Committee (29). London, H. M. Stationery Office.

St. Louis Report (25). Mental Testing in St. Louis. Public School Messenger, Vol. 23, No. 1.

Terman, L. M. (25). *Mental and Physical Traits of a Thousand Gifted Children.* Genetic Studies of Genius, Vol. I. Stanford University Press.

Thorndike, E. L. et al. (26). "On the Variability of Boys and Girls from Thirteen to Eighteen." *Ped. Sem.*, 33, 182-184.

Thorndike, E. L. et al. (26). "Sex Differences in Status and Gain in Intelligence Scores from Thirteen to Eighteen." *Ped. Sem.*, 33, 167-181.

Wallin, J. E. W. (30). "Differences in Chronological Age, Mental Capacity, and Sex Ratios of Children Referred from Many School Systems as Candidates for Special Classes." *J. of Appl. Psych.*, 14, 1-31.

Whipple, G. M. (27). "Sex Differences in Intelligence Test Scores in the Elementary School." *J. of Ed. Res.*, 15, 111-117.

Whipple, G. M. (27). "Sex Differences in Army Alpha Scores in the Secondary School." *J. of Ed. Res.*, 15, 269-275.

Winsor, A. L. (27). "The Relative Variability of Boys and Girls." *J. of Ed. Psych.*, 18, 327-336.

CHAPTER XXIV

THE INHERITANCE OF INTELLIGENCE

The Pioneer Work of Galton.—That all sorts of traits, physical as well as mental, are inherited is a truism at the present time. The work of the biologist has now gone far beyond establishing the fact of inheritance. He is investigating and throwing light upon the mechanism of inheritance. The biologist has for the most part been concerned with physical characteristics. The laws governing the transmission of such traits are equally true for mental traits.

The pioneer work in the study of the inheritance of mental characteristics was done by Galton (69, 74, 83), the founder of the eugenics movement. He begins his book, "Hereditary Genius," with this significant sentence, "I propose to show in this book that a man's natural abilities are derived by inheritance, under exactly the same limitations as are the form and physical features of the whole organic world." His work is a study of the family history of 977 eminent men. He finds that these men had 332 eminent fathers, brothers or sons and 203 eminent grandfathers, grandsons, uncles or nephews, that is, a total of 535 eminent relatives. The probable number of eminent relatives for 977 average men is estimated at 4. Other studies by Galton deal with the inheritance of scientific aptitude, artistic ability, as well as with such physical traits as stature, eye color and certain diseases. Further studies of eminent men were later made by Cattell (03) and Ellis (04).

We may say that Galton's work contained the elements of two methods for the study of inheritance, namely, the family history method, and the correlational method. The family history method is a detailed study of all the members of a par-

ticular family group, in order to show the repeated occurrence of a given trait in the family group. If the trait is present in several generations, the chances are that it is inherited. The correlational method, on the other hand, concerns itself with showing that related individuals are more alike with reference to certain traits than are unrelated individuals. This resemblance is generally expressed by a coefficient of correlation. We shall survey the evidence for the inheritance of intelligence as offered by these two methods.

The Family History Method.—Apart from Galton's work dealing with the family histories of eminent British men, the first extensive study by the family history method was made by Dugdale (77) in America. It is doubtful whether this work was due to the influence of Galton. The work, however, is important, because it later stimulated many other studies in this country. Dugdale shows how the "Jukes" family, descended from Max Jukes in the eighteenth century, increased and multiplied, always exhibiting the three traits of prostitution, crime and pauperism. Dugdale tends to speak of these three factors as if they were directly inherited. We would be inclined now to say that the descendants of Max Jukes inherited certain mental, moral and physical traits which would lead to crime, prostitution and pauperism in the ordinary environment of civilized life.

It is of interest to note here the continuance of the Jukes family in the twentieth century. Estabrook (16) went over Dugdale's work and brought the family history down to 1915. While Dugdale's genealogy includes 709 persons, Estabrook's contains 2,820. Of this number Estabrook estimates that 131 or 4½ per cent were so feebleminded as to need custodial care. Of the 1,258 descendants living in 1915 he estimates that 110 were mentally defective, 83 intemperate and 171 industrious. Some members of the family had emigrated to a different section of the country, but the changed environment seemed to lead to no improvement so long as members of the old stock intermarried. Improvement of the old stock took place by

out-breeding into better stocks. Estabrook feels that the outstanding characteristic of the family is feeblemindedness, although many individuals do not show that degree of mental defect which at present is supposed to require custodial care.

This work of Dugdale stimulated many other workers, so that we have to-day numerous family-history studies showing the inheritance of all sorts of mental and physical characteristics. For our purpose we need only mention some of the studies showing the inheritance of intelligence or lack of intelligence. Goddard's (14) study of the Kallikak Family is one of the most interesting of these. Beginning with a feebleminded girl in an institution for mental defectives the family tree was traced back to the time of the Revolution. Here we find the illegitimate mating of a normal man with a feebleminded woman leading to a long line of feebleminded individuals. A later, legitimate mating of the same man with a woman of superior intelligence leads to a long line of highly intelligent and more or less eminent individuals. Goddard (14) also presents the family histories of 327 feebleminded cases. The evidence here presented for the inheritance of feeblemindedness is very impressive. As Goddard says, "One hundred sixty-four or 54 per cent of the remaining 300 histories show other feebleminded persons in such numbers or in such relation to the individual case studied as to leave no doubt of the hereditary character of the mental defect. In these cases it is evident from the charts themselves that we are dealing with a condition of mind or brain which is transmitted as regularly and surely as color of hair or eyes. Thirty-four cases, 11.3 per cent, have been grouped under the head of Probably Hereditary. The charts of these, while not showing so certainly as in the former group the hereditary nature of the trouble, yet have a high degree of probability and may be considered hereditary."

The type of family-history study represented by Goddard's work has been repeated by numerous workers. The "Win" tribe, an Indian-negro-white mixture, has been described by Estabrook and McDougle (26). The authors assume that the

original white stock was normal and that the level of intelligence was lowered by the mixing of negro and Indian blood.

In a study of one thousand cases of young repeated offenders Spaulding and Healy (14) show the inheritance of mental and physical defects, but find no evidence of the inheritance of criminalistic traits as such. The mental make-up of families containing congenitally deaf children was studied by Pintner and Osborn (19). Altogether 520 individuals were included. As many individuals as possible were given intelligence tests and the results of such tests revealed a large percentage of poor intelligence.

The Correlational Method.—The use of the coefficient of correlation to express the amount of resemblance in any trait among related individuals offers a simple and valuable measure for the study of heredity. If the coefficient is low and approaches zero, there is no resemblance in the trait in question among the individuals compared. Chance pairs of unrelated individuals show zero correlations for psychological traits, because there is no factor at work making for a resemblance among such chance selections of individuals. If, now, there is a greater resemblance among related than among unrelated individuals, this will show itself in the higher correlation among the related individuals. And we find such higher correlations among related individuals. This does not constitute a proof that heredity is the factor causing the resemblance, but, since the two groups are the same in every respect except in the matter of relationship or non-relationship, it would seem that the factor of heredity is the cause of the resemblance.

Pearson (04) gives the following correlations for certain physical traits:

Color of eyes	brothers	.52
Height	brothers	.50
Height	father and son	.30
Cephalic index	brothers	.49
Color of hair	brothers	.55

For certain mental characteristics we have the following correlations:

	Brothers	Sisters	Brother and Sister
Vivacity	.47	.43	.49
Self-assertiveness	.53	.44	.52
Introspection	.59	.47	.63
Popularity	.50	.57	.49
Conscientiousness	.59	.64	.63
Temper	.51	.49	.51
Ability	.46	.47	.44
Handwriting	.53	.56	.48
Average	.52	.51	.52

These are samples of Pearson's correlations. He finds the means for physical characteristics and for mental characteristics to be about the same, namely, .50, and hence he concludes that mental characteristics are inherited in the same way as are physical characteristics.

With reference to deafness, Schuster (05) finds the resemblance between fathers and children to vary from .46 to .62, and between mothers and children from .45 to .62. Earle (03) found the correlation between children of the same family for spelling to be .50. Starch (17) gave educational and psychological tests to 18 pairs of adult siblings who were university students. The average of 15 correlations for the educational tests is .42. The four psychological tests give an average coefficient of .38. "The resemblance of siblings is apparently no greater in those mental traits which are directly affected by school work than in those which are not so affected. This seems to indicate that the mental similarities of children of the same parents are due primarily to heredity rather than to similarity of environment, since the resemblance is no greater in those traits which are more directly affected by environment."

Pintner (18), using a combination of six mental tests, found the following coefficients of association (Yule's Q) for siblings and non-siblings:

	Q	No. of Pairs
School A	47	91
School B	28	89
Both schools	39	180
Chance selection of non-siblings	14	151
Another chance selection of non-siblings	19	300

The Pearson coefficient for 180 pairs of siblings is .22.

The results of the correlational method with more modern tests of general intelligence have been summarized in Tables 31 and 32. The first table shows us the results for twins where the resemblance is high; the next table for siblings, where the resemblance is less.

TABLE 31

THE RESEMBLANCE BETWEEN TWINS

Author	*r*	*n*	*Tests*	*Comments*
Thorndike (05)	.69 to .90	50	Various	Siblings .29 to .32
	.83	..	Various	Younger Twins
	.70	..	Various	Older twins
Merriman (24)	.81	47	Binet	Ages 5–9
	.76	58	Binet	Ages 10–16
	.78	28	Beta	Ages 5–9
	.66	48	Beta	Ages 10–16
	.80	54	N. I. T.	Ages 5–10
	.87	89	N. I. T.	Ages 11–18
Wingfield (28)	.75	102	N. I. T. & Multi-Mental	All pairs
	.70	57	N. I. T. & Multi-Mental	Fraternal twins
	.90	45	N. I. T. & Multi-Mental	Identical twins
Holzinger (29)	.88	50	Binet I.Q.	Identical twins
	.92	50	Otis I.Q.	Identical twins
	.63	52	Binet I.Q.	Fraternal twins
	.62	52	Otis I.Q.	Fraternal twins

If we glance at the coefficients of correlation in Table 31 we note that they range from .62 to .92. All of the workers divide the twins into older and younger and they find no tendency for the coefficient to rise with the older twins due to a longer period of similar environment. Wingfield's division into fraternal and identical shows an extremely high coefficient for identical twins. In general we may think of a coefficient of .75 as representing the resemblance between twins.

Let us turn now to Table 32 which summarizes the results for siblings. Here our coefficients range from .27 to .68. The central tendency of these coefficients is about .50. Wherever the author reports comparisons with non-sibling groups, we find the coefficients drop down to zero. Hildreth (25) tested the children in an orphanage and found a correlation of .32 for 253 pairs of sibs. For 78 pairs of sibs reared apart the correlation was .24, which is somewhat less than .32 for all sibs, but still not nearly zero. In addition, she found a correlation of —.10 for unrelated children reared in the same environment (the orphan home) for 10 to 25 per cent of their lives, and a correlation of —.17 for unrelated children reared in the same environment for 50 to 100 per cent of their lives. If sibs are reared apart, they do resemble each other in intelligence; if non-sibs are reared together, they do not resemble each other in intelligence. Thorndike's (28) study deals with high school children only. He concludes it with this paragraph: "If we may accept Pearson's results for the resemblance of siblings in eye color, hair color, and cephalic index, and regard .52 ± .016 as the resemblance in traits entirely free from environmental influence, we may infer that the influence upon intelligence of such similarity in environment as is caused by being siblings two to four years apart in age in an American family to-day is to raise the correlation from .52 to .60."

Finally we have one study of resemblance beyond the first degree of relation. Dexter (24) finds a correlation of .22 for 131 pairs of cousins tested by means of intelligence tests.

TABLE 32

THE RESEMBLANCE BETWEEN SIBLINGS

Author	*Sibs*		*Non-Sibs*		*Tests*	*Comments*
	r	*n*	*r*	*n*		
Gordon (19)	.53	91	..	..	Binet	Orphan Children
Gordon (18–20)	.61	216	..	..	Binet	Orphans
Elderton (23)	.53	216	..	..	Binet	Gordon's data
	.27 to .67	Various	..	..	Various	
Hart (24)	.45	252	..	..	Various	
	.46	147	..	..	Binet	Rural
	.40	219	..	..	Binet	Urban
Madsen (24)	.63	63	−.04	63	Binet	
Pintner (24)	.33	53	+.09	53	Group Intelligence	
Merriam (24)	.49	784	..	..	Various	Average of 7 r's
Hildreth (25)	.68	1,028	−.19	100	Binet	
Thorndike (28)	.60	Large	..	..	I. E. R.	High School only

If now we sum up the results of these correlational studies, we find we have a hierarchy of coefficients like this:—

Identical twins	.90
All twins	.75
Fraternal twins	.70
Siblings	.50
Cousins	.20
Unrelated individuals	.00

And this is exactly the hierarchy we should expect to find, if heredity were an important factor in intelligence. At the same time these findings are no proof of the inheritance of intelligence, because this hierarchy of coefficients corresponds also to the degree of similarity of environment usually found among the various groups of individuals. All that we can claim is that our correlations agree perfectly with the hypothesis that intelligence is inherited in the same way as physical characteristics are inherited.

Other Results.—There are other studies which are neither family history studies nor correlational studies. Some of these are studies of pairs of twins separated at an early age. If intelligence is inherited and if twins are very much alike in intelligence, then they should be similar in intelligence, even though separated at an early age and brought up in different circumstances. Three pairs of twins have been studied by Newman (29) and one by Muller (25). In all these cases the twins were reared apart, Muller's cases being separated at two weeks old and not meeting until 18 years old. Two pairs differed by 12 points in I.Q. In each case the twin having the higher I.Q. was brought up in the better environment. In another case one twin had a year more of formal education and showed a difference of plus three months on the Binet Scale and plus 16 months on the Stanford Achievement Test, thus showing a much greater gain in education than in intelligence. The last pair both scored very high in intelligence within a few points of each other on two group intelligence tests. All of

these four pairs were identical twins. The greatest divergence reported is 12 points in I.Q. This is not great. In no case did environmental influences make one twin superior and the other dull, i.e., cause a difference of 30 or 40 points in I.Q. All of the differences reported might readily occur with a correlation of .9 for identical twins. We must remember that there is no necessity for identical twins to test exactly alike. There is some suggestion that better education produces a slight increase in I.Q., but on the whole these four cases fit well into the theory of general intelligence being largely inherited. We are able to detect this in spite of our rough measures of intelligence.

Cobb and Hollingworth (25) studied the siblings of a group of children of high I.Q. This group was especially chosen for their very high I.Q.s, the average being 154. The average of the siblings was 129. The sibs of those having I.Q.s above 150 had an average I.Q. of 133, for those below 150 the average was 124. As would be anticipated, there is a marked regression of the sibs toward the mean of the general population, but "those who test in the best 0.5 per cent of intelligence have siblings all within the upper 50 per cent." If we pick children of very high I.Q., naturally their siblings will be lower, since the correlation between siblings is only .5. So Moorrees (24) started with children of very low I.Q. in an institution, median I.Q. 29, and found that the median I.Q. of the parents was 63.

A somewhat different type of study has been made by Hirsch (28). He studied 1,945 East Kentucky mountaineers. He found an average I.Q. of 79 on group intelligence tests. He believes this low mentality is due to selective migration and close inbreeding over a long period of time. He presents evidence to show the migration of the better types and the numerous cousin marriages among those who have been left behind.

Social Status.—Indirectly the factor of heredity can be seen at work in the differences in intelligence found among children of different social status. In the long run those possessing superior intelligence will in general tend to occupy the higher

types of positions in the world, and those possessing inferior intelligence will gravitate towards the lower occupations. Since children tend to inherit the same kind of intelligence as their parents, we ought to find differences in the intelligence of children as we proceed from the lower to the higher occupations of their parents. The results of many workers show this to be the case.

Table 33 attempts to summarize several reports. The classification of occupations differs somewhat from one report to another, hence the many blank spaces in our table. Furthermore, the reader must note that some of the reports are in terms of median or average I.Q., others in per cents above the median and still others in raw score. Hence the actual numbers in the table are to be disregarded. The main thing to be noted is the trend of the values as we run down from the top to the bottom of each column. This trend is obviously from high to low as we go down the columns. Now this trend is present at the earliest ages in the pre-school child and persists through elementary, high school and college. Our college data are the least satisfactory and are based on very few cases, and of the three reports two are from the same college. The general picture, however, is clear. The I.Q. of the child corresponds to the parental occupation, and this is as true of very young children as of older children. The Haggerty and Nash (24) report includes both high school and elementary children and a comparison between the I.Q.s of these groups is, therefore, possible. We note the selective influence of the high school by the higher median I.Q.s for all occupations, and the narrower range of I.Q.s. Furthermore, children of unskilled laborers are not now at the bottom. Relatively few of these go to high school and then only the brighter ones. The value of the pre-school results is to show that these differences in I.Q.s of occupational groups are not caused by education. They exist in marked degree even before the child comes to school.

Two reports by English workers differ so much in occupa-

TABLE 33

Intelligence of Children by Parental Occupation

Occupation	Pre-School		Elementary School					High School		High School and College	College		
	Goodenough (28) Mean I.Q.	*Goodenough (28) Test II Mean I.Q.*	*Bridges (17) C. M. A. on Point Scale*	*Pressey (19) % above Group Median*	*Dexter (23) Mean I.Q.*	*Collins (28) Mean I.Q.*	*Haggerty and Nash (24) Median I.Q.*	*Haggerty and Nash (24) Median I.Q.*	*Book (22) % above State Median*	*Sandiford (26) Median I.Q.*	*Bear (26) Median score*	*Bear (28) Median score*	*West (29) Mean I.Q.*
Professional	116	125	142	85	115	114	116	121	60	105	51	50	11
Semi-Professional	112	120											
Managers						113							
Executive				68					54			46	
Clerical	108	113			106	112			60	103			
Business					104		107	112			48	50	109
Salesmen			126						56		52	56	
Proprietors			121										
Tradesmen						109							
Foremen						106							
Skilled			112		99	102	98	108		102.5	43	53	
Artisan				41					55				
Farmers						99	91	106	43	102	37	43	111
Semi-Skilled	105	108			92		95	108		102			
Slightly Skilled	104	107											
Unskilled	96	96	83		89	94	89	111		101			
Laborers				39					47				105
Unknown						96					35	45	
n	380	380	300	548	2,782	4,727	8,121	8,121	—	5,052	95	172	?

tional classification that they could not be summarized in Table 33. They are, therefore, presented separately in Table 34. The Duff and Thomson (23) report includes all children of ages 11 and 12 in the county of Northumberland, excluding Newcastle. The Macdonald (25-26) report deals with children of ages 11 to 13 in the Isle of Wight. The similarity of the two reports is striking, and the general picture is the same as that in Table 33 for the United States and Canada.

TABLE 34

INTELLIGENCE OF CHILDREN BY PARENTAL OCCUPATION (ENGLAND)

Occupation	*Duff and Thomson (23) Mean I.Q.*	*Macdonald (25-26) Mean I.Q.*
Professional	112.2	107
Managers	110.0	109
Higher commercial	109.3	103
Army, navy, police, post.	105.5	100
Shopkeepers	105.0	101
Engineers	102.9	101
Foremen	102.7	103
Ship builders	....	99
Building trades	102.0	99
Metal workers	100.9	...
Miscellaneous industrial	100.6	99
Miners, quarrymen	97.6	98
Seamen	....	97
Agriculture	97.6	97
Laborers	96.0	96
n	13,419	2,047

We have stressed the fact of the importance of the Goodenough (28) report for pre-school children. Gesell and Lord (27) confirm these results by their comparison of eleven pairs of nursery school children, one of each pair being from a wealthy and the other from a poor home. We still need a study based upon infants to round out this work.

Goodenough (27) has attacked the problem from a slightly different angle by asking whether the education of the mother, who spends more time with the child, correlates higher with the child's intelligence than does the education of the father. Her cases number 380 children from 18 to 54 months. She finds no difference between the correlations of the two parents and concludes that the results of the test are largely determined by heredity.

Another way of showing the relationship between the intelligence of the child and the social status of the family is by means of correlations. We have four such reports:

	r	*n*	
Chapman and Wiggins (25)	+.32	632	I.Q. on N. I. T.
Stroud (28)	+.25	1,057	Tax assessment and I.Q.
Chauncey (29)	+.21	113	Grade VIII. Sims' score.
	+.19	130	Grade IX. Sims' score.
Sirkin (29)	+.36	845	Grade IV
	+.43	850	Grade V
	+.39	682	Grade VI

These correlations are low, but positive. The measures of social status are probably not very accurate. Stroud (28) uses objective data in his tax assessments, but notes the limited range, which reduces his correlation. The Sirkin (29) data are for Russian children. Taken as a whole the results fit into the hypothesis of inherited intelligence.

Terman's (25) gifted children had a much larger percentage of fathers in higher occupational groups as compared with the general population from which they were chosen:

	Gifted Per Cent	*General Per Cent*
Professional	29.1	2.9
Public Service	4.5	3.3
Commercial	46.2	36.1
Industrial	20.2	57.7

Duff's (29) comparison of an intelligent group (I.Q. above 135) with a control group (I.Q. 100) showed the following percentages of parental occupations:

	Intelligent Per Cent	*Control Per Cent*
Industry—higher	6	0
Professional	13	0
Industry—lower	27	12
Clerical	9	8
Skilled	9	8
Total higher	64	28
Semi-skilled	30	50
Unskilled	6	23
Total lower	36	73

Similarly Allen (26) investigating 49 families each of which contained a gifted child with I.Q. above 132 found 70 per cent of the fathers in the highest group of Taussig's classification. And Jones and Carr-Saunders (26-27) find that the I.Q. of orphan children correlates with parental occupation just as it does with children not in orphanages, although the orphan child has largely the same school and "home" experiences.

We may sum up these studies of the relationship of the intelligence of the child to the occupational or social status of the parent by saying that they fit in very well with the theory that intellectual potentiality is largely inherited. As we have seen in a previous chapter, the I.Q. is not wholly impervious to environmental influence, yet on the whole it seems to measure roughly that potentiality for intellectual development which seems to be inherited.

We must remember, of course, that the occupational status of an individual is by no means a sure guide to his mentality. It is only in a general sense that occupational status correlates with mental ability. The distribution of children in all occu-

pational groups runs from very low to very high intelligence. Since it is the duty of education to make the most of all the mental ability of all the pupils, educational classification according to social status is not justified. The brighter children in the lower social groups should be given just as much opportunity as they can profit by. The road through high school and college should be open to all who have the intelligence and the interest to travel along it.

Race Mixtures.—The results of the inheritance of general intelligence seem also to appear in a general way in tests of mulattoes and pure negroes. If colored subjects are classified according to skin color, the intelligence of the lighter negroes is superior to that of the darker negroes. The greater the mixture of white blood, the higher the intelligence of the group becomes. Thus Ferguson (21), summarizing the army results, says, "a summary of the test results indicates that roughly 20 per cent of the pure negroes, 25 per cent of the negroes three-fourths pure, 30 per cent of the true mulattoes and 35 per cent of the quadroons equal or exceed the average score of comparable whites." A rougher classification into two groups, lighter and darker, gives the same results. "The lighter class contained those whose color indicated that they were true mulattoes or persons of a larger proportion of white blood than true mulattoes. The darker class contained pure negroes and those whose skin color indicated that they had a smaller proportion of white blood than true mulattoes. The classification was made by the various examiners of the groups. In Alpha the lighter negroes obtained a median score of 50; the darker obtained a median of 30. In Beta, the lighter negroes obtained a median score of 36; the darker obtained a median of 29." Similar facts are also shown by Ferguson (16) with reference to negro children in Virginia when tested by several mental tests. When classified into four groups according to skin color, he found that the average scores on the tests increased from the darkest up to the lightest groups.

Garth (21, 22) has shown the general superiority of mixed

to full blood Indians on several mental tests. The mixed blood group is 11 per cent better than the full blood group in performing tests of higher mental processes. The difference in performance between the mixed and full bloods does not seem to be so marked or so clearly defined as is the case with the light and dark negroes. Garth is not sure whether we have here a racial difference or merely a difference due to nurture.

The Indian has also been studied by Hunter (22), who shows that the ability involved in the Otis test decreases with a decrease in the amount of white blood. A gradual decrease in ability from the quarter to the full blood is shown on the intelligence test scores as follows:

	¼ Blood	*½ Blood*	*¾ Blood*	*Full Blood*
25 percentile ...	77	68	56	36
Median	109	91	78	67
75 percentile ...	128	118	108	94

In the case of the negro and perhaps in the case of the Indian we have a race of inferior intelligence as measured by our present intelligence tests when compared with American whites. The greater the amount of white blood entering into the various mixtures of the two races the greater is the intelligence of the resulting progeny, and this takes place because of the inheritance of mental ability.

Summary.—Although the belief in the inheritance of mental ability may be old, it is only during the last century that we have accumulated scientific evidence of importance. This evidence takes the form of family-history investigations and measures of amount of resemblance between related individuals. All sorts of mental traits as well as that complex of traits called general intelligence are inherited in the same way as purely physical traits. The potency of environment is not nearly so great as is commonly supposed. Intelligence tests have done much to show that all children are not created free and equal with respect to their mental abilities. A child's abilities are determined by his ancestors, and all that environ-

ment can do is to give opportunity for the development of his potentialities. It cannot create new powers or additional abilities. This, then, is the main function of education, to measure the inherited capacities of the child and to so arrange the environment as to give full opportunity for all these capacities to develop to the uttermost.

REFERENCES

Allen, G. (26). "The Families Whence High Intelligence Springs." Eugenics Record Office, *Bull. No. 25.*

Bear, R. M. (26). "Factors in the Achievement of College Freshmen." *Sch. and Soc.,* 24, 802-804.

Bear, R. M. (28). "Factors Affecting the Success of College Freshmen." *J. of Appl. Psych.,* 12, 517-523.

Book, W. F. (22). *The Intelligence of High School Seniors.* Chap. X. New York.

Bridges, J. W., and Coler, L. E. (17). "The Relation of Intelligence to Social Status." *Psych. Rev.,* 24, 1-31.

Cattell, J. McK. (03). "A Statistical Study of Eminent Men." *Pop. Sci. Monthly,* 62, 359-377.

Chapman, J. C., and Wiggins, D. M. (25). "Relation of Family Size to Intelligence of Offspring and Socio-Economic Status of Family." *Ped. Sem.,* 32, 414-421.

Chauncey, M. R. (29). "The Relation of the Home Factor to Achievement and Intelligence Test Scores." *J. of Ed. Res.,* 20, 88-90.

Cobb, M. V., and Hollingworth, L. S. (25). "The Regression of Siblings of Children Who Test at or Above 135 I.Q. (Stanford-Binet)." *J. of Ed. Psych.,* 16, 1-7.

Collings, J. E. (28). "The Intelligence of School Children and Parental Occupation." *J. of Ed. Res.,* 17, 157-169.

Davenport, C. B. (13). *Heredity in Relation to Eugenics.* New York.

Dexter, E. S. (23). "The Relation Between Occupation of Parent and Intelligence of Children." *Sch. and Soc.,* 17, 612-614.

Dexter, E. S. (24). "On Family Resemblance Beyond the First Degree of Relation." *Sch. and Soc.,* 19, 501-502.

Duff, J. F. (29). "Children of High Intelligence." *Brit. J. of Psych.,* 19, Part 4, 413-438.

Duff, J. F., and Thomson, G. H. (23). "The Social and Geographical Distribution of Intelligence in Northumberland." *Brit. J. of Psych.,* 14, Part 2, 192-198.

Dugdale, R. L. (77). *The Jukes.* Fourth Edition. 1910.

Earle, E. L. (03). "The Inheritance of the Ability to Learn to Spell." Columbia Contribs. to Phil., Psych. and Ed., Vol. II, No. 2, 41-44.

Elderton, E. M. (23). "A Summary of the Present Position with Regard to the Inheritance of Intelligence." *Biometrika*, 14, 378-408.

Ellis, H. (04). *A Study of British Genius.* Hurst and Blackett, London.

Estabrook, A. H. (16). *The Jukes in 1915.* Carnegie Institute of Washington.

Estabrook, A. H., and McDougle, I. E. (26). *Mongrel Virginians.* Baltimore.

Ferguson, G. O. (16). "The Psychology of the Negro." *Arch. of Psych.*, No. 36.

Ferguson, G. O. (21). "The Mental Status of the American Negro." *Sci. Monthly*, 12, 533-543.

Galton, F. (69). *Hereditary Genius.* New York, 1880. (Reprint.)

Galton, F. (74). *English Men of Science.* New York, 1890. (Reprint.)

Galton, F. (83). *Inquiries into the Human Faculty and Its Development.* New York.

Garth, T. M. (21). "The Results of Some Tests on Full and Mixed Blood Indians." *J. of Appl. Psych.*, 5, 359-372.

Garth, T. M. (22). "A Comparison of Mental Abilities of Mixed and Full Blood Indians on a Basis of Education." *Psych. Rev.*, 29, 221-236.

Gesell, A., and Lord, E. E. (27). "A Psychological Comparison of Nursery School Children from Homes of Low and High Economic Status." *Ped. Sem.*, 34, 339-356.

Goddard, H. H. (14). *Feeblemindedness, Its Causes and Consequences.* New York.

Goddard, H. H. (14). *The Kallikak Family.* New York.

Goodenough, F. L. (27). "The Relation of the Intelligence of Pre-School Children to the Education of their Parents." *Sch. and Soc.*, 26, 54-56.

Goodenough, F. L. (28). "The Relation of the Intelligence of Pre-School Children to the Occupation of Their Fathers." *Am. J. of Psych.*, 40, 284-294.

Gordon, K. (18-20). "The Influence of Heredity on Mental Ability." Report of Children's Dept., State Board of Control of California.

Gordon, K. (19). "Psychological Tests of Orphan Children." *J. of Delinq.*, 18, 46-55.

Haggerty, M. E., and Nash, H. B. (24). "Mental Capacity of Children and Parental Occupation." *J. of Ed. Psych.*, 15, 559-572.

Hart, H. (24). "Correlations Between Intelligence Quotients of Siblings." *Sch. and Soc.*, 20, 382.

Hildreth, G. H. (25). *The Resemblance of Siblings in Intelligence and Achievement.* T. C. Contribs. to Educ., No. 186.

Hirsch, N. D. M. (28). "An Experimental Study of the East Kentucky Mountaineers." *Gen. Psych. Monog.*, 3, 189-245.

Holzinger, K. J. (29). "The Relative Effect of Nature and Nurture Influences on Twin Differences." *J. of Ed. Psych.*, 20, 241-248.

Hunter, W. S. (22). "The Relation of Degree of Indian Blood to Score on the Otis Intelligence Test." *J. of Comp. Psych.*, 2, 257-277.

Jones, D. C., and Carr-Saunders, A.M. (26-27). "The Relation Between Intelligence and Social Status Among Orphan Children." *Brit. J. of Psych.*, 17, 343-364.

Macdonald, H. (25-6). "The Social Distribution of Intelligence in the Isle of Wight." *Brit. J. of Psych.*, 16, 123-129.

Madsen, I. N. (24). "Some Results with the Stanford Revision of the Binet-Simon Tests." *Sch. and Soc.*, 19, 559-562.

Memoirs (21). The National Academy of Sciences, Vol. XV, Chap. 8. Intelligence of the Negro.

Merriam, C. (24). "Intellectual Resemblance of Twins." *Psych. Rev. Monog.*, 152, 1-58.

Moorrees, V. (24). "The Immediate Heredity of Primary Aments Committed to a Public Institution." *J. of Appl. Psych.* 8, 89-127.

Muller, H. J. (25). "Mental Traits and Heredity." *J. of Hered.*, 16, 433-448.

Newman, H. H. (29). "Mental and Physical Traits of Identical Twins Reared Apart." *J. of Hered.*, 20, 49-64, 97-104, 153-166.

Pearson, K. (02-3). "On the Laws of Inheritance in Man." *Biometrika*, 2, 356-462.

Pearson, K. (04). "On the Laws of Inheritance in Man." *Biometrika*, 3, 131-190.

Pintner, R. (18). "The Mental Indices of Siblings." *Psych. Rev.*, 25, 252-255.

Pintner, R. (24). "Training Students in Group Intelligence Testing." *J. of Ed. Res.*, 9, 271-280.

Pintner, R., and Osborn D. (19). "The Mentality of the Families of the Congenitally Deaf." *Am. Annals of the Deaf*, Vol. 64, 96-134.

Pressey, S. L., and Ralston, R. (19). "The Relation of the General Intelligence of School Children to the Occupation of Their Fathers." *J. of Appl. Psych.*, 3, 366-373.

Sandiford, P. (26). "Paternal Occupations and Intelligence of Offspring." *Sch. and Soc.*, 23, 117-119.

Schuster, E. (05-06). "Hereditary Deafness." *Biometrika*, 4, 465-482.

Sirkin, M. (29). "The Relation Between Intelligence Age and Home Environment of Elementary School Pupils." *Sch. and Soc.*, 30, 304-308.

Spaulding, E. R., and Healy, W. (14). "Inheritance as a Factor in Criminality. A Study of a Thousand Cases of Young Repeated Offenders." *Bull. of the Am. Acad. of Medicine*, 15, 1-24.

Starch, D. (17). "The Similarity of Brothers and Sisters in Mental Traits." *Psych. Rev.*, 24, 235-238.

Stroud, J. B. (28). "A Study of the Relation of Intelligence Test Scores of Public School Children to the Economic Status of Their Parents." *J. of Genet. Psych.*, 35, 105-111.

Terman, L. M. et al. (25). *Genetic Studies of Genius*. Stanford University Press.

Thorndike, E. L. (05). "Measurements of Twins." *Arch. of Phil., Psych. and Scientific Methods,* No. 1.

Thorndike, E. L. (28). "The Resemblance of Siblings in Intelligence." *Twenty-seventh Yearbook.* Nat'l Soc. for the Study of Educ., Part I, Chap. III, 41-53.

West, L. S. (29). "A Preliminary Analysis of the Student Body of Battle Creek College." *Eugenics,* 2, 4, 23-31.

Wingfield, A. H. (28). *Twins and Orphans.* The Inheritance of Intelligence. London, Dent.

Wingfield, A. H., and Sandiford, P. (28). "Twins and Orphans." *J. of Ed. Psych.* 19, 410-423.

CHAPTER XXV

MISCELLANEOUS

Intelligence tests are now being used in so many investigations that we have accumulated a mass of information on various relationships between intelligence and other factors. The topics of this chapter do not fit readily into any other chapter of the book. They are not logically related, and hence this chapter will be scrappy and fragmentary in nature, but it will give the reader an idea as to how widespread the use of intelligence tests has been. In no sense is this chapter meant to be exhaustive. We shall merely attempt to give indications of the type of work that has been done and refer to a few of the workers.

Physical Characteristics.*—There seems to be practically no correlation between height or weight and intelligence when C.A. is kept constant. The correlations are all low (Brooks, 28; Murdoch and Sullivan, 23). McHale (26) found no difference in I.Q. or M.A. between comparable under-, over-, and normal-weight children. Height-weight ratios or morphological indices correlate very low or zero with intelligence (Heidbreder, 26; Sheldon, 26-27; Garrett and Kellogg, 28). In general there are very low correlations between intelligence and all kinds of physical tests (Gates, 24) and this is true of cranial capacity (Estabrooks, 28), of ossification and dentition (Abernethy, 25), of pigmentation of hair and eyes (Estabrooks, 29), and of age of pubescence (Viteles, 29). Perkins (26) is one of the few workers who finds significant correlations between dentition and intelligence. Gordon (23) finds the left-handed

* See the recent book by D. G. Paterson, *Physique and Intellect,* New York, 1930, for a thorough discussion of this topic.

slightly higher in I.Q. than the right-handed but Haefner (28) finds no difference in intelligence between the right-handed and the left-handed. In general, therefore, we may say that these physical characteristics correlate slightly, if at all, with general intelligence.

With reference to the gross motor development, Cunningham (27) finds substantial correlations with mental age for infants and young children. Hertzberg (29) finds positive correlations, but points out that motor dexterity alone is of no value for predicting M.A. Among college students, athletes and non-athletes are about the same in intelligence (Hindman, 29), and correlations with specific athletic performances, e.g., hundred yard dash, broad jump, etc., are zero (Landis et al. 23).

With reference to the common physical defects of school children, Mallory (22) finds that intelligence is not associated with most of these, with the exception of nasal obstruction ($Q = .10$) and hearing defects ($Q = .36$). Westenberger (27), however, finds that the common minor defects of school children seem to have no influence upon intelligence or school achievement. The removal of adenoids and tonsils has no effect on the I.Q. (Rogers, 22; Lowe, 23-24), at least for a year after the operation. The average number of hookworms increases as the I.Q. decreases, but there is no evidence for an increase in I.Q. after treatment (Smillie and Spencer, 26). The mean I.Q. of 49 crippled children was found to be 84, and that of their 89 sibs to be 89 according to Fernald and Arlitt (25). The mean I.Q. of 194 crippled was 82, the spastic birth paralysis cases being lowest with a mean of 69 for 27 cases. Stutterers do not differ from non-stutterers according to the results of many intelligence tests given by McDowell (28). Fox (28) does not find that glandular therapy leads to any gain in I.Q., nor can Woods (26) detect any gain in I.Q. from teaching and practicing nutrition in special classes for the dull, although considerable gains in educational achievement resulted. De Weerdt (28) reports the case of a boy whose I.Q. changed from

116 to 130 during the period June, 1925, to December, 1927, due to improvement in physical condition. In general, however, we may sum up by saying that there is only a slight correlation between physical defect and intelligence, the correlation varying greatly with the type of physical defect. Ordinary treatment of school children for minor physical defects does not seem to influence the I.Q.

Character and Emotional Traits.—Intelligence is negatively correlated with cheating to the extent of about —.5 to —.6 according to the thorough studies of Hartshorne and May (28). If the home background is kept constant this drops to about —.4. In general, therefore, children of high intelligence are likely to be more honest than children of low intelligence. Tests of service and intelligence correlate +.16 and Hartshorne and May (29) remark: "In all populations the dull children are definitely less helpful and charitable." However, the brightest are not any more helpful than those of average I.Q. Tests of persistence correlate +.15 with intelligence and tests of self-control practically zero.

There are innumerable studies of tests of non-intellectual characteristics and many of them give the correlations of such traits with various measures of intelligence. We cannot here describe all of these studies. Table 35 gives samples of these

TABLE 35

Correlation of Intelligence and Other Traits

Trait	*r with Intelligence*	*Author*
Cheating	—.50 to —.60	Hartshorne and May (28)
Honesty in examinations (university students)	slightly positive	Fenton (27)
Service	+.16	Hartshorne and May (29)
Cooperation	+.14	Maller (29)
Persistence	+.15	Hartshorne and May (29)

Trait	*r with Intelligence*	*Author*
Self-control	about zero	Hartshorne and May (29)
Moral knowledge	+ .69	Hartshorne and May (27)
Judgment of good and bad	+ .24	Woodrow (26)
Happiness	— .05	Watson (30)
Overstatement	+ .29	Woodrow and Bemmels (28)
Aggressiveness	+ .008	Gilliland (26)
Suggestibility	— .75	Otis (24)
Caution	+ .40	Brown (23-24)
Liberalism	+ .28	Symonds (25)
Social perception	+ .10	Gates (25)
Consequences	+ .49	Chassell et al. (24)
Will-temperament	+ .20 to + .77	Downey (23)
Will-temperament	+ .60	Meier (23)
Will-temperament	+ .35	Traxler (25)
Will-temperament	+ .08	Uhrbrock (28)
Will-temperament (non-verbal)	— .23 to + .39	Uhrbrock (27)
Emotionality (Pressey X-O)	+ .04 to + .17	Thompson and Remmers (28)
Introversion-extroversion	+ .02	Hoitsma (25)
Introversion-extroversion	zero	Guthrie (27-28)
Introversion-extroversion	low	Hovey (29)
Neurotic inventory	+ .04	Thurstone and Thurstone (30)

various studies. If we study the table we note that, in general, intelligence would seem to be positively correlated with desirable traits. Sometimes the correlation is zero, but in no case do we see any marked positive correlation between intelligence and an undesirable trait. Tests of emotionality, introversion-extroversion, and neurotic tendencies seem to have no correlation with intelligence. Character testing is only beginning and

many of these tests have low reliability and validity, but the general picture of the positive correlation of desirable traits is probably a true one.

There remain a few studies, the results of which cannot be expressed in terms of correlations. Hurlock (24) found that both praise and reproof raised the I.Q. five points above the control group in re-tests of 408 children. The superior children were stimulated more than the inferior by praise and reproof. McClure (29) asked teachers to list their problem children. Out of an enrolment of 26,364, he obtained 533 cases. The mean I.Q. of 499 of these cases was 90.1, so problem cases would seem to be slightly below average in intelligence. Brownell (28) finds that college cribbers are slightly below the average on intelligence tests, and Cummings (27-28) finds that deserters in the navy are lower in I.Q. than recruits in general.

Judging Intelligence.—The estimation of intelligence by means of physical signs seems very uncertain. Using photos, Pintner (18) found practically no correlation, but Gaskill et al. (27) obtained a median rank correlation of $+.42$. Moriwaki (29) obtained a correlation of $+.28$ by averaging the judgments of four judges in the case of 15 photos of normal school students. Omwake's (25) correlations for 30 judges ranking the photos of 30 men and 30 women range from $-.15$ to $+.33$. For handwriting the same author reports correlations from $-.16$ to $+.16$. Inflection of the voice in reading a passage when rated by experts correlates from $+.34$ to $+.50$ with intelligence according to Michael and Crawford (27).

Evidently photos and handwriting can be of no practical value for the measurement of intelligence. Interviews may be a little better, for Moriwaki's (29) correlation of .28 for photos rises to .56 for an interview. But Magson (26) can only get a correlation of .15 for five minute interviews and objective tests of intelligence, and he concludes that it is impossible to estimate the intelligence of an unknown individual by means of a short interview.

Size of Family.—There is a slight negative correlation between intelligence and size of family. Several investigators have found this to be the case:

Author	*r*	*Comments*
Clark (22-23)	—.08	Delinquent boys
Bradford (25)	—.25	Ages 10-11
Sutherland and Thomson (26-27)	—.20	1924 cases—ages 10½ to 11½
Lentz (27)	—.30	4,330 cases
Willoughby (28)	—.30	Fertility of college stocks
Thurstone and Jenkins (29)	—.09	10,000 clinic cases
Sutherland (29)	—.13	One occupational group

These data would seem to point to a correlation of —.2 to —.3 for I.Q. and size of family. Bradford believes this to mean a 50 per cent increase in the group of dull children in the next two generations provided no radical change takes place. Lentz finds a steady decrease in I.Q. from families of one (I.Q. 107.9) to families of 12 or more (I.Q. 79.9). The significance of Sutherland's data is that all his cases are drawn from one occupational group, namely, miners, and yet the negative correlation is still present. This negative correlation between I.Q. and size of family indicates probably a gradual decrease in intelligence in the population at large, unless it is off-set by a higher mortality rate for those of lower I.Q. We must also take into account the length of time during which such a negative correlation may persist.

Birth Order.—The few studies on the relation of intelligence to birth order show conflicting results. Thurstone and Jenkins (29) find that the mean intelligence rises with order of birth, that is, the later born have the higher I.Q.s. Jones and Hsiao (28) find no difference between the difference in intelligence of adjacent and non-adjacent pairs of siblings and hence they conclude that there is no difference in birth order.

McFadden (29) finds that the I.Q.s tend to decrease from the youngest to the oldest child. Arthur (26) finds no difference between younger and older siblings for 70 pairs of American-born children, but a difference of 6.8 points of I.Q. in favor of the younger sibling for 92 pairs of children of foreign parents. From such conflicting results we can draw no general conclusion, but must wait for more thorough and extensive studies.

Month of Birth.—Blonsky (29) presents data showing the mean I.Q. for 453 cases according to month of birth. The lowest mean I.Q. is for the winter months and the highest for the spring. From this he argues that the effects of sunlight, air and better food in spring during early life affect the I.Q. favorably. Pintner (31), however, gives the mean I.Q.s for 4,925 cases according to month of birth and can find no reliable differences between any two months or any of the seasons. He concludes that the effects claimed by Blonsky are not apparent in the United States for a random sampling of children. At present, therefore, this suggested influence of the month of birth on I.Q. is doubtful.

Homogamy in Intelligence.—A few studies of the degree of correspondence in intelligence between husband and wife have been made. Jones (29) has discussed the available data and we reproduce his summary here, together with one added report by Worcester (23).

Test	*r*	*P.E.*	*Corrected r*	*Investigator*
5 verbal tests	.40	.06	.44	Willoughby
6 non-verbal	.37	.06	.44	Willoughby
Stanford-Binet . . .	.42	.04	.47	Burks
Stanford-Binet . . .	.55	.05	.62	Burks
Otis S.A.	.49	.04	.55-.60	Freeman et al.
Army Alpha	.60	.04	.61	Jones
Army Alpha	.66	.04	. . .	Worcester

Bi-lingualism.—All reports seem to show the disadvantages of bi-lingualism. Smith (22-23) gave many tests to monoglot and bi-lingual Welsh children and found a general superiority

for the monoglots. He believes that bi-lingualism tends to mental confusion. Saer (23-24) also tested Welsh children and found the bi-linguals inferior. He regards this as due to confusion caused by the conflicts of two languages at an early age. Hauck (29) in Germany finds upper Silesian children below the intelligence of middle or south German children and believes bi-lingualism is one cause for this. He supposes that bi-lingualism exercises an inhibiting influence upon mental development. Jamieson and Sandiford (28) find monoglot Indians superior to bi-lingual Indians on all tests except the Pintner-Paterson Performance Test. All these studies point in the same direction, namely the inferiority of the bi-lingual individual. They are merely suggestive and by no means conclusive. Whether bi-linguals are inferior in all types of intelligence or only in verbal tests should be further investigated. The whole problem raised by these studies is an interesting one. It impinges upon the question of different kinds of intelligence; it also touches upon educational policies as to the best age for beginning the study of a foreign language.

Intelligence and Other Abilities.—General intelligence of the verbal type correlates low with tests of mechanical abilities. Thus Bell (24) reports correlations of + .02 to + .24 for men, and — .01 to + .13 for women students between verbal intelligence tests and the Stenquist Mechanical Tests. Kefauver (28-29) finds a higher correlation for boys between I.Q. and shop work, namely + .22. Schulz (28) finds no correlation between intelligence and motor capacity as tested by the Viteles Machine Feeding Test.

Intelligence and musical talent correlate only slightly; + .24 according to Brown (28) and from + .01 to + .32 for the separate Seashore tests according to Fracker and Howard (28).

The Lewerenz Art Test correlates only + .15 with intelligence according to Lewerenz (28).

Some maze problems make good intelligence tests according to Cox (28). Students of high intelligence made fewer errors and used the trial and error method less frequently. Speed in

solving the maze correlates from + .09 to + .59 with intelligence. Crosland et al. (29) found no relation between intelligence and amount of error on the Müller-Lyer illusion, nor did intelligence correlate with improvement resulting from practice. Training university students in methods of study is valuable for those above the 25th percentile in intelligence, but those below this point do not profit sufficiently to pass in their work (Pressey, 28). The greatest gains in intellect are made by those of high intelligence regardless of what high school studies they may take during the year (Thorndike, 23).

Leaders in different activities in school come in general from the upper three quarters in intelligence, according to Caldwell and Wellman (26), who do not find a high correlation between intelligence and leadership among a rather select group of children. Henig (27) believes that a negative correlation exists between intelligence and number of accidents among boys in a vocational school.

Play and Chums.—Lehman and Witty (27) find that the dull take more to social games, the gifted more to sedentary and solitary games, and in general C.A. is more potent than M.A. in influencing a boy's play behavior (Lehman and Wilkerson, 28). Furfey (27) studied 62 pairs of chums and found the correlation for C.A. to be .39, while that for M.A. was only .24, but Warner (23) maintains that boys' gangs come closer together in M.A. than in C.A.

School Factors.—Among high school students there is a negative correlation between intelligence and time spent in study (— .01 to — .08), but positive between I.Q. and out-of-school educational activities and between I.Q. and extra-curricular activities (Uhl, 28). Gains in intelligence tests continue during the school vacation, whereas gains in educational tests are zero or negative (Nelson, 28).

Speed.—Many workers have reported correlations between various time limits on intelligence tests. Allowing more time in most cases does not materially alter the relative positions of those examined (Ruch, 24; Walters, 27; Freeman, 28). Peak

and Boring (26) believe that time is of great import in intelligence tests and that speed of reaction probably correlates highly with intelligence. But Farnsworth (27) finds that simple reaction time does not correlate with intelligence test scores, although choice reaction correlates from + .14 to + .53. Similarly Sisk (26) finds a correlation of — .014 between Army Alpha and simple speed tests, and + .23 between Army Alpha and complex speed tests. Considering the various aspects of intelligence, Clark (25) finds that speed correlates lowest with intelligence + .54, whereas level correlates + .70 and range + .77. And Hunsicker (25) finds that rate of work at zero level of difficulty is positively correlated with level of intelligence. From all these results we may deduce the practical conclusion that our time limits on group intelligence tests are in all probability quite fair to most workers, but in general it is advisable so far as possible to have fairly generous time-limits and place more emphasis on difficulty or level.

Work Conditions.—Farnsworth (28) could find no difference in intelligence tests scores, if his subjects worked alone or in groups. Wyatt (26-27) found that monotonous work (soap wrapping) had only a slight positive correlation with intelligence, namely, + .25.

Other Indicators of Intelligence.—Conrad and Jones (29) believe that fidelity of report in the form of a questionnaire based on motion pictures would make a good intelligence test. They found correlations of .68 and .71 between scores on Army Alpha and such a questionnaire given to individuals ranging in age from 10 to 54. McGeoch (28) finds correlations of .27 to .44 for children in ability to report.

Knowledge of Intelligence Rating.—Allen (30) failed to find any ill effects resulting from giving students their scores on intelligence tests. Fenton (24) found that 108 students were not affected, 9 were encouraged and 8 discouraged according to their own statements as a result of knowing their scores.

Examiners.—Pintner (24) believes that thorough training for both individual and group intelligence testing is necessary.

Rogers (28) finds that the correlation between first test and re-test is lower for student examiners than for experts. Marine (29) does not find that familiarity of the examiner with the child influences the reliability of the examination when carried out by an expert examiner.

Scoring Intelligence Tests.—Pintner (26) finds many errors made in scoring such group tests as the N.I.T., and hence the need for practice in scoring tests. He believes that tests should be constructed with more objective scoring methods. Dearborn and Smith (29) report that 73 per cent of test blanks re-scored were found to contain errors. The major cause of error was due to differences in subjective estimate, where such estimate had to be made by the scorer. Pressey (26) has devised an apparatus for scoring multiple choice items. From such reports as these it is obvious that there is still much to be done in making intelligence tests that can be more accurately and rapidly scored.

Conclusions.—No attempt will be made to summarize this chapter, for it is in a sense a brief summary in itself of much of the miscellaneous work in which intelligence testing has figured. It shows how widespread intelligence testing has become and how much interest has centered around the relationship between intelligence and other factors. At the same time it will be obvious to the reader that many problems have been touched on in this chapter that are still awaiting solution. The future will undoubtedly see more work in many of the problems that have been raised here.

REFERENCES

Abernethy, E. M. (25). "Correlations in Physical and Mental Growth." *J. of Ed. Psych.*, 16, 458-466.

Allen, C. M. (30). *Some Effects Produced in an Individual by Knowledge of His Own Intellectual Level.* T. C. Contribs. to Educ. No. 401.

Arthur, G. (26). "The Relation of I.Q. to Position in Family." *J. of Ed. Psych.*, 17, 541-550.

Bell, J. C. (24). "Mechanical Aptitude and Intelligence." *Contribs. to Ed.,* Vol. I. N. Y. Soc. for the Exper. Study of Ed., 270-282.
Blonsky, P. P. (29). *"Früh und Spätjahrkinder." Jahrbuch für Kinderheilkunde,* 124, 115-120.
Bradford, E. J. G. (25). "Can Present Scholastic Standards be Maintained." *Forum of Educ.,* 3, 186-198.
Brooks, F. D. (28). "The Organization of Mental and Physical Traits During Adolescence." *J. of Appl. Psych.,* 12, 228-241.
Brown, A. W. (28). "The Reliability and Validity of the Seashore Tests of Musical Talent." *J. of Appl. Psych.,* 12, 468-476.
Brown, W. M. (23). "Character Traits as Factors in Intelligence Test Performance." *Arch. of Psych.,* No. 65.
Brown, W. M. (24). "A Study of the Caution Factor." *Am. J. Psych.,* 35, 368-386.
Brownell, H. C. (28). "Mental Test Traits of College Cribbers." *Sch. and Soc.,* 27, 764.
Caldwell, O. W., and Wellman, B. (26). "Characteristics of School Leaders." *J. of Ed. Res.,* 14, 1-13.
Chassell, C. F., et al. (24). "A Test of Ability to Weigh Foreseen Consequences." *T. C. Record,* 25, 39-50.
Clark, J. R. (25). *The Relation of Speed, Range and Level to Scores on Intelligence Tests.* Ph. D. Dissertation, Columbia University, New York.
Clark, W. W. (22-23). "Birth Rate and Native Intelligence." *Psych. Clin.,* 14, 111-115.
Conrad, H. S., and Jones, H. E. (29). "Fidelity of Report as a Measure of Adult Intelligence." *Univ. of Calif. Pub. in Psychol.,* 3, 245-276.
Cox, C. M. (28). "The Intelligence Factor in the Solution of Space Problems with the Two-Storey Maze." *Am. J. of Psych.,* 40, 542-561.
Crosland, H. R., Taylor, H. R., and Newsom, S. J. (29). "Practice and Improvability in the Müller-Lyer Illusion in Relation to Intelligence." *J. of Gen. Psych.,* 2, 290-306.
Cummings, D. E. (27-28). "Iron Men for Iron Ships." *Person, J.,* 6, 87-92.
Cunningham, B. V. (27). "An Experiment in Measuring Gross Motor Development of Infants and Young Children." *J. of Ed. Psych.,* 18, 458-464.
Dearborn, W. F., and Smith, C. W. (29). "The Results of Rescoring Five Hundred Thirty Dearborn Tests." *J. of Ed. Psych.,* 20, 177-183.
De Weerdt, E. H. (28). "The Effect of Malnutrition upon Expression of Intelligence." *Sch. and Soc.,* 27, 540-541.
Downey, J. E. (23). *The Will-Temperament and Its Testing.* Yonkers.
Estabrooks, G. H. (28). "The Relation between Cranial Capacity, Relative Cranial Capacity and Intelligence in School Children." *J. of Appl. Psych.,* 12, 524-529.

Estabrooks, G. H. (29). "Intelligence and Pigmentation of Hair and Eyes in Elementary School Children." *Am. J. of Psych.*, 41, 106-108.

Farnsworth, P. R. (28). "Concerning So-called Group Effects." *J. of Genet. Psych.*, 35, 587-594.

Farnsworth, P. R. et al. (27). "Speed in Simple and Serial Action as Related to Performance in Certain Intelligence Tests." *Ped. Sem.*, 34, 537-551.

Fenton, N. (24). "Mental Test Scores and Self-Regard." *Ed. Admin. and Super.*, 10, 103-108.

Fenton, N. (27). "An Objective Study of Student Honesty During Examinations." *Sch. and Soc.*, 26, 341-344.

Fernald, M. R., and Arlitt, A. H. (25). "A Psychological Study of a Group of Crippled Children of Various Types." *Sch. and Soc.*, 21, 449-452.

Fox, E. J. (28). "An Investigation of the Effect of Glandular Therapy on the Intelligence Quotient." *Ment. Hyg.*, 12, 90-102.

Fracker, G. C., and Howard, V. M. (28). "Correlation Between Intelligence and Musical Talent Among University Students." *Psych. Monog.*, 39, No. 178, 157-161.

Freeman, F. S. (28). "Power and Speed: Their Influence upon Intelligence Test Scores." *J. of Appl. Psych.*, 12, 631-635.

Furfey, P. H. (27). "Some Factors Influencing the Selection of Boys' Chums." *J. of Appl. Psych.*, 11, 47-51.

Gates, A. I. (24). "The Nature and Educational Significance of Physical Status and of Mental, Physiological, Social and Emotional Maturity." *J. of Ed. Psych.*, 15, 329-358.

Gates, G. S. (25). "A Preliminary Study of a Test for Social Perception." *J. of Ed. Psych.*, 16, 452-457.

Garrett, H. E., and Kellogg, W. N. (28). "The Relation of Physical Constitution to General Intelligence, Social Intelligence and Emotional Instability." *J. of Exp. Psych.*, 11, 113-129.

Gaskill, P. C., Fenton, N., and Porter, J. P. (27). "Judging the Intelligence of Boys from Their Photographs." *J. of Appl. Psych.*, 11, 394-403.

Gilliland, A. R. (26). "A Revision and Some Results with the Moore-Gilliland Aggressiveness Test." *J. of Appl. Psych.*, 10, 143-150.

Gordon, K. (23). "Some Notes on the Mental Status of the Left-handed." *J. of Delinq.*, 8, 154-157.

Guthrie E. R. (27-28). "Measuring Introversion and Extroversion." *J. of Ab. and Soc. Psych.*, 22, 82-88.

Haefner, R. (28). *The Educational Significance of Left-handedness.* T. C. Contribs. to Ed., No. 360.

Hartshorne, H., May, M. A., et al. (27). *Testing the Knowledge of Right and Wrong.* The Religious Ed. Assoc. Monog. No. 1.

Hartshorne, H., and May, M. A. (28). *Studies in the Nature of Character.* Vol. I, Studies in Deceit. New York.

Hartshorne, H., and May, M. A. (29). *Studies in the Nature of Character.* Vol. II, Studies in Service and Self-Control. New York.

Hauck, E. (29). *Zur differentiellen Psychologie des Industrie und Landkindes.* Langensalza, Beltz.

Heidbreder, E. (26). "Intelligence and the Height-Weight Ratio." *J. of Appl. Psych.,* 10, 52-62.

Henig, M. S. (27). "Intelligence and Safety." *J. of Ed. Res.,* 16, 81-87.

Hertzberg, O. E. (29). "The Relationship of Motor Ability to the Intelligence of Kindergarten Children." *J. of Ed. Psych.,* 20, 507-519.

Hindman, D. A. (29). "Athletics and Scholarship at the Ohio State University." *Sch. and Soc.,* 30, 93-96.

Hoitsma, R. K. (25). "The Reliability and Relationships of the Colgate Mental Hygiene Test." *J. of Appl. Psych.,* 9, 293-303.

Hovey, H. B. (29). "Measures of Extraversion-Introversion Tendencies and their Relation to Performance Under Distraction." *J. of Genet. Psych.,* 36, 319-329.

Hunsicker, L. M. (25). *A Study of the Relationship Between Rate and Ability.* T. C. Contribs. to Ed., No. 185.

Hurlock, E. B. (24). "The Value of Praise and Reproof as Incentives for Children." *Arch. of Psych.,* No. 71.

Jamieson, E., and Sandiford, P. (28). "The Mental Capacity of Southern Ontario Indians." *J. of Ed. Psych.,* 19, 313-328.

Jones, H. E. (29). "Homogamy in Intellectual Abilities." *Am. J. of Soc.,* 35, No. 2, September.

Jones, H. E., and Hsiao, H. H. (28). " A Preliminary Study of Intelligence as a Function of Birth Order." *J. of Genet. Psych.,* 35, 428-433.

Kefauver, G. N. (28-29). "Relationship of the Intelligence Quotient and Scores on Mechanical Tests with Success in Industrial Subjects." *Vocational Guid. Mag.,* 7, 198-207.

Landis, M. H., Burtt, H. E., and Nichols, J. H. (23). "The Relation Between Physical Efficiency and Intelligence." *Am. Phys. Ed. Rev.,* May, 1-2.

Lehman, H. C., and Wilkerson, D. A. (28). "The Influence of Chronological Age versus Mental Age on Play Behavior." *J. of Genet. Psych.,* 35, 312-321.

Lehman, H. C., and Witty, P. A. (27). *The Psychology of Play Activities.* New York.

Lentz, T. (27). "Relation of I.Q. to Size of Family." *J. of Ed. Psych.,* 18, 486-496.

Lewerenz, A. S. (28). "I.Q. and Ability in Art." *Sch. and Soc., 27,* 489-490.

Lowe, G. M. (23-24). "Mental Changes After Removing Tonsils and Adenoids." *Psych. Clin.,* 15, 92-100.

Magson, E. H. (26). "How We Judge Intelligence." *Brit. J. of Psych.* Mono. Suppl., No. 9.

Maller, J. B. (29). *Cooperation and Competition.* T. C. Contribs. to Ed., No. 384.

Mallory, J. N. (22). *A Study of the Relation of Some Physical Defects to Achievement in the Elementary School.* Geo. Peabody Coll. Contrib. to Educ., No. 9.
Marine, E. L. (29). *The Effect of Familiarity with the Examiner upon Stanford-Binet Test Performance.* T. C. Contribs. to Ed. No. 381.
McClure, W. E. (29). "Characteristics of Problem Children Based on Judgments of Teachers." *J. of Juv. Res.,* 13, 124-140.
McDowell, E. D. (28). *Educational and Emotional Adjustments of Stuttering Children.* T. C. Bureau of Pub., N. Y.
McFadden, J. H. (29). "A Further Note on the Differential I.Q.'s of Siblings." *J. of Appl. Psych.,* 13, 86-91.
McGeoch, J. A. (28). "Intelligence and the Ability to Report." *Am. J. of Psych.,* 40, 596-599.
McHale, K. (26). *Comparative Psychology and Hygiene of the Overweight Child.* T. C. Contribs. to Ed., No. 221.
Meier, N. C. (23). "A Study of the Downey Test by the Method of Estimates." *J. of Ed. Psych.,* 14, 385-395.
Michael, W., and Crawford, C. C. (27). "An Experiment in Judging Intelligence by the Voice." *J. of Ed. Psych.,* 18, 107-114.
Moriwaki, E. (29). "Note on the Comparative Validities of Judgments of Intelligence Based on Photographs and on Interviews." *J. of Appl. Psych.,* 13, 630-631.
Murdoch, K., and Sullivan, L. R. (23). "A Contribution to the Study of Mental and Physical Measurements in Normal Children." *Am. Phys. Ed. Rev.,* May and June.
Nelson, M. J. (28). "The Abilities and Achievements of Elementary School Pupils Before and After Vacation." *Sch. and Soc.,* 28, 371-372.
Omwake, K. T. (25). "The Value of Photographs and Handwriting in Estimating Intelligence." *Public Personnel Studies,* 3, 2-15.
Otis, M. (24). "A Study of Suggestibility of Children." *Arch. of Psych.,* No. 70.
Paterson, D. G. (30). *Physique and Intellect.* New York.
Peak, H., and Boring, E. G. (26). "The Factor of Speed in Intelligence." *J. of Exp. Psych.,* 9, 71-94.
Perkins, F. J. (26). "The Relation of Dentition to Mental Age." *Ped. Sem.,* 33, 387-398.
Pressey, L. C. (28). "The Permanent Effects of Training in Methods of Study on College Success." *Sch. and Soc.,* 28, 403-404.
Pressey, S. L. (26). "A Simple Apparatus Which Gives Tests and Scores and Teaches." *Sch. and Soc.,* 23, 373-376.
Pintner, R. (18). "Intelligence as Estimated from Photographs." *Psych. Rev.,* 25, 286-296.
Pintner, R. (24). "Training Students in Group Intelligence Testing." *J. of Ed. Res.,* 9, 271-280.
Pintner, R. (26). "Accuracy in Scoring Group Intelligence Tests." *J. of Ed. Psych.,* 17, 470-475.
Pintner, R. (31). *Intelligence and Month of Birth.* To appear shortly.

Rogers, A. L. et al. (28). "The Constancy of the I. Q. and the Training of Examiners." *J. of Ed. Psych.*, 19, 257-262.
Rogers, M. C. (22). "Adenoids and Diseased Tonsils. Their Effect on General Intelligence." *Arch. of Psych.*, No. 50.
Ruch, G. M. (24). "The Speed Factor in Mental Measurements." *J. of Ed. Res.*, 9, 39-45.
Saer, D. J. (23-24). "The Effect of Bilingualism on Intelligence." *Brit. J. of Psych.*, 14, 25-38.
Schultz, R. S. (28). "A Test for Motor Capacity in the Industries and in School." *J. of Appl. Psych.*, 12, 169-189.
Sheldon, W. H. (26-27). "Morphologic Types and Mental Ability." *J. of Person. Res.*, 5, 447-451.
Sisk, T. K. (26). *The Interrelations of Speed in Simple and Complex Responses.* Peabody College for Teachers. Contribs. to Educ., No. 23.
Smillie, W. G., and Spencer, C. R. (26). "Mental Retardation in School Children Infested with Hookworms." *J. of Ed. Psych.*, 17, 314-321.
Smith, F. (22-23). "Bilingualism and Mental Development." *Brit. J. of Psych.*, 13, 271-282.
Sutherland, H. E. G., and Thomson, G. H. (26-27). "The Correlation Between Intelligence and Size of Family." *Brit. J. of Psych.*, 17, 81-92.
Sutherland, H. E. G. (29). "The Relationship Between I.Q. and Size of Family." *J. of Ed. Psych.*, 20, 81-90.
Symonds, P. M. (25). "A Social Attitudes Questionnaire." *J. of Ed. Psych.*, 16, 316-322.
Thompson, L. A., and Remmers, H. H. (28). "Some Observations Concerning the Reliability of the Pressey X-O Test." *J. of Appl. Psych.*, 12, 477-494.
Thorndike, E. L. (23). "Mental Discipline in High School Studies." *J. of Ed. Psych.*, 15, 1-22 and 83-98.
Thurstone, L. L., and Jenkins, R. L. (29). "Birth Order and Intelligence." *J. of Ed. Psych.*, 20, 641-651.
Thurstone, L. L., and Thurstone, T. G. (30). "A Neurotic Inventory." *J. of Soc. Psych.*, 1, 3-30.
Traxler A. E. (25). "The Will Temperament of Upper-Grade and High School Students." *School Rev.*, 33, 264-273.
Uhl, W. L. (28). "The Relation of Intelligence to Time Expenditures In and Out of School." *Ed. Admin. and Super.*, 14, 179-184.
Uhrbrock, R. S., and Downey, J. S. (27). "A Non-verbal Will-Temperament Test." *J. of Appl. Psych.*, 11, 95-105.
Uhrbrock, R. S. (28). *An Analysis of the Downey Will-Temperament Tests.* T. C. Contribs. to Educ., No. 296.
Viteles, M. S. (29). "The Influence of Age of Pubescence upon the Physical and Mental Status of Normal School Students." *J. of Ed. Psych.*, 20, 360-368.
Walters, F. C. (27). *A Statistical Study of Certain Aspects of the Time Factor in Intelligence.* T. C. Contribs. to Educ., No. 248.

Warner, M. L. (23). "Influence of Mental Level in the Formation of Boys' Gangs." *J. of Appl. Psych.*, 7, 224-236.

Watson, G. (30). "Happiness Among Adult Students of Education." *J. of Ed. Psych.*, 21, 79-109.

Westenberger, E. J. (27). "A Study of the Influence of Physical Defects upon Intelligence and Achievement." *Cath. Univ. Am. Ed. Res. Bull.*, 2, No. 9.

Willoughby, R. R. (28). "The Survival of Intelligence." *Proceedings Nat. Acad. Sci.*, 14, 892-894.

Worcester, D. A. (23). "Does Like Like Like?" *Sch. and Soc.*, 18, 655.

Woodrow, H. (26). "A Picture-Preference Character Test." *J. of Ed. Psych.*, 17, 519-531.

Woodrow, H., and Bemmels, V. (28). "Overstatement in Third-Grade Children." *J. of Appl. Psych.* 12, 404-416.

Woods, E. L. (26). "Intellectual and Emotional Improvement in Nutrition Classes." *Ed. Res. Bull.*, Los Angeles Schools, Vol. 6, No. 1, 6-11.

Wyatt, S. (26-27). "An Experimental Study of a Repetitive Process." *Brit. J. of Psych.*, 17, 192-209.

INDEX OF SUBJECTS

INDEX OF NAMES

I

J

K

Date Due			